OM of: US-08-711-417C-165 to: SPTREMBL_19:*   out_format : pfs

Date: Aug 28, 2002 10:16 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+.n2p.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/6228611/rrunat_28082002_100211_13637/app_query.fasta_1.1639
 -DB=SPTREMBL_19 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=6228611_@CGN1_1_383 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-08-711-417C-165
 Query length: 1551
 Database: SPTREMBL_19:*
 Database sequences: 562222
 Database length: 172994929
 Search time (sec): 172.940000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
sp_vertebrate:Q9PUG1	+	1964.00	2593.41	8.6e-137	482	! Q9pug1 ambystoma mexicanum (axo
sp_vertebrate:Q918W1	+	1810.00	2388.82	2.1e-125	500	! Q918w1 raja eglanteria. ikaros.
sp_vertebrate:Q90XA3	+	1762.50	2325.32	6.7e-122	526	! Q90xa3 brachydanio rerio (zebra
sp_vertebrate:Q93581	+	1747.50	2304.56	9.4e-121	537	! Q93581 brachydanio rerio (zebra
sp_vertebrate:Q918V9	+	1402.50	1848.14	2.6e-95	513	! Q918v9 raja eglanteria. aiolos
sp_vertebrate:Q918W2	+	1396.00	1839.35	7.9e-95	522	! Q918w2 raja eglanteria. helios.
sp_vertebrate:Q90XA2	+	1373.50	1811.34	3.5e-93	434	! Q90xa2 brachydanio rerio (zebra
sp_vertebrate:Q90WJ6	+	1351.50	1782.30	1.5e-91	428	! Q90wj6 seriola quinqueradiata (
sp_vertebrate:Q90W82	+	1341.50	1766.32	8.6e-91	563	! Q90w82 homo sapiens (human). zi
sp_human:Q9UKT9	+	1303.50	1716.93	5.4e-88	509	! Q9ukt9 homo sapiens (human). ai
sp_rodent:O08900	+	1302.00	1714.98	6.9e-88	507	! O08900 mus musculus (mouse). ai
sp_human:Q96JP3	+	1296.50	1706.97	1.8e-87	545	! Q96jp3 homo sapiens (human). ki
sp_rodent:Q9Z2Z2	+	1273.50	1676.68	8.9e-86	533	! Q9z2z2 mus musculus (mouse). eo
sp_vertebrate:Q9PU55	+	1260.00	1659.53	8.7e-85	494	! Q9pu55 gallus gallus (chicken).
sp_human:Q9H2S9	+	1252.50	1649.81	3.1e-84	483	! Q9h2s9 homo sapiens (human). zi
sp_vertebrate:Q90XA1	+	1220.50	1609.44	6.8e-82	392	! Q90xa1 brachydanio rerio (zebra
sp_vertebrate:Q90X58	+	1109.50	1460.24	1.1e-73	481	! Q90x58 ambystoma mexicanum (axo
sp_vertebrate:Q9I8U5	+	1007.50	1324.41	3.9e-66	507	! Q9i8u5 petromyzon marinus (sea
sp_vertebrate:P79751	+	906.00	1311.76	2.4e-65	417	! P79751 fugu rubripes (japanese
sp_vertebrate:O42244	+	806.00	1061.50	2.7e-51	328	! O42244 xenopus laevis (african
sp_vertebrate:Q90XA0	+	693.00	912.72	5.8e-43	294	! Q90xa0 brachydanio rerio (zebra
sp_human:Q96LD7	+	678.00	892.50	7.5e-42	304	! Q96ld7 homo sapiens (human). he
sp_vertebrate:Q9I8W0	+	676.50	893.89	8.8e-42	216	! Q9i8w0 raja eglanteria. eos (fi
sp_vertebrate:O13100	+	672.00	892.36	1.7e-41	138	! O13100 xenopus laevis/gilli. dr
sp_human:Q96NN3	+	448.50	593.32	5.8e-25	180	! Q96nn3 homo sapiens (human). co
sp_human:Q9H5V7	+	395.50	514.67	6.0e-21	419	! Q9h5v7 homo sapiens (human). co
sp_human:Q9H2T0	+	394.50	513.34	7.1e-21	419	! Q9h2t0 homo sapiens (human). zi
sp_invertebrate:Q95US1	+	365.50	467.93	1.2e-18	846	! Q95us1 locusta migratoria (mi
sp_invertebrate:Q95US0	+	349.00	451.02	1.7e-17	512	! Q95us0 locusta migratoria (mi
sp_human:Q9H9P0	+	341.50	442.09	6.0e-17	462	! Q9h9p0 homo sapiens (human). co
sp_human:Q9H7R5	+	333.50	428.68	2.5e-16	613	! Q9h7r5 homo sapiens (human). th
sp_invertebrate:Q95UR8	+	332.50	424.35	3.3e-16	830	! Q95ur8 schistocerca americana
sp_human:Q9Y2P0	+	331.50	426.94	3.5e-16	559	! Q9y2p0 homo sapiens (human). bc
sp_invertebrate:Q9N9D4	+	331.50	425.92	3.6e-16	620	! Q9n9d4 megaselia abdita. puta
sp_human:Q9NYW8	+	330.00	423.19	4.4e-16	714	! Q9nyw8 homo sapiens (human). rb
sp_human:Q96SE7	+	330.00	420.93	5.0e-16	839	! Q96se7 homo sapiens (human). zi
sp_human:Q9P0L1	+	324.50	414.70	1.2e-15	754	! Q9p0l1 homo sapiens (human). zi
sp_invertebrate:Q95UR9	+	324.00	418.34	1.4e-15	948	! Q95ur9 schistocerca americana
sp_human:Q96JH5	+	324.00	411.77	1.4e-15	488	! Q96jh5 homo sapiens (human). ki
sp_rodent:O70162	+	322.50	411.29	1.8e-15	814	! O70162 mus musculus (mouse). my

sp_human:Q9H6Z6	+	321.50	417.17	1.7e-15	393	! Q9h6z6 homo sapiens (human).
sp_human:Q96NJ6	+	320.00	412.75	2.4e-15	502	! Q96nj6 homo sapiens (human).
sp_human:Q9Y2A4	+	320.00	409.88	2.6e-15	671	! Q9y2a4 homo sapiens (human).
sp_human:O75850	+	317.00	408.74	4.0e-15	504	! O75850 homo sapiens (human).
sp_human:Q9P2J8	+	315.00	402.12	6.2e-15	752	! Q9p2j8 homo sapiens (human).

seq_name: sp_vertebrate:Q9PUG1

seq_documentation_block:
ID Q9PUG1   PRELIMINARY;   PRT;   482 AA.
AC Q9PUG1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TRANSCRIPTION FACTOR IKAROS.
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20093994; PubMed=10630298;
RA Durand C., Charlemagne J., Fellah J.S.;
RT "Structure and developmental expression of Ikaros in the Mexican
RT axolotl.";
RL Immunogenetics 50:336-343(1999).
DR EMBL; AF186359; AAF01038.1; -.
DR InterPro; IPR000822; znf-C2H2.
DR Pfam; PF00096; zf-C2H2; 4.
DR SMART; SM00355; ZnF_C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 482 AA; 53740 MW; 57351794 9D590F08 CRC64;

alignment_scores:
              Quality: 1964.00   Length: 517
                Ratio: 4.454   Gaps: 5
Percent Similarity: 85.300   Percent Identity: 72.534

alignment_block:
US-08-711-417C-165 x Q9PUG1

Align seg 1/1 to: Q9PUG1 from: 1 to: 482

    1 ATGGATGCTGACGAGGTCAAGACATGCTTTCTCATCAGGGAGGAAAG 50
      |||||||||||||||||||||||||::|||||||||:::::|||||||
    1 MetAspAlaAspGluAlaGlnGlnMetSerGlnValSerGlyAsnGlnSe 17

   51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
      |||||||||||:::|||||||||||||||||||||:|||||||||||||:|
   17 rSerProMetSerAspIleProGluAspGlyAspGluProMetProValP 34

  101 CCGAGGACCTCTCCACCACCTCGGGAGGACGAGCAAAGTCCAAGAGTGAC 150
      |||:|||||||:|||||::|||||||||:|||||:||||||||:||||||
   34 roGluAspLeuSerThrAsnSerGlyAlaGlnHisHisPheArgAsnGlu 50

  151 AGAGTCGTGGCCAGTAATGTTAAACTAGAGACTCAGAGTGATGAAGAGAA 200
      ||||||::|||||||||||||||||||||||:|||||::|||:::||||||
   51 LysAlaLeuAlaSerAsnValLysLeuGluAlaGlnLeuSerAspGluAs 67

  201 TGGGCTGCCTGTGAAATGAAATGAAATGGGGAAGAAATGTCCGAGGATTTACGAA 250
      |||:||:|||::||||||||:|||||||||||||||:||||||||||||||||||
   67 nGlyHisAlaCysAspMetAsnGlyGluGlyLysAlaGluAspLeuArgM 84

  251 TGCTTGATGCCTCGGGAGGAGAAAAATGAATGGCTCCCACCAGGGACCAAGGC 300
      ||||||||||||||:||||||||||||||||:|||||:|||||||||||||:|||
   84 etLeuAspAlaSerGlyGluLysMetAsnGlyLeuHisAspGlyProGly 100

  301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGCGAAAACT 350
      |||:::|||||:||||||||||||||||||||

```
101 SerLysAlaLeuSerGlyValGly...................... 108

351 AAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400

108 ............... 108

401 TTCCAAAAGAAGCCACACTGGAGAACGGCCCCTTCCAGTGCAATCAGTGC 450
109 .........GlyGluArgProPheGlnCysAsnGlnCys 118

451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTTCCGGCACATCAAGCTGCA 500
119 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 135

501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGGCCTGCCGCC 550
135 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 152

551 GGAGGACGGCCCTCACTGGCCACCTGAGGACGGACCACTCCGTTGTAAACCT 600
152 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 168

601 CACAAATGTGATATTTGTGGCGAAGCTATAAAACGAGCGTCTTTAGA 650
169 HisLysCysGlyTyrCysGlyGluArgSerTyrLysGlnArgSerLeuGl 185

651 GGAACATAAAGACGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
185 uGluHisLysGluArgCysHisSerAsnTyrLeuGlnSerMetGlyLeuProG 202

701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
202 lyAsnLeuTyrAlaValVallLysGluGluThrAsnProSerAspMetPro 218

751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTGCTGGACAGACT 800
219 GluAspLeuCysLysIleGlySerGluArgSerLeuMetLeuAspArgLe 235

801 AGCAAGTAATGTCGCCAAACGTAAGACGTAAGACTCCTATGCCTCAGAAATTTCTTG 850
235 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheValG 252

851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACG...TACGAG 897
252 lyGluLysArgLeuSerAspHisProPheAspAlaSerThrProTyrGlu 268

898 AAGGAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAA 947
269 LysGluAsnGluMetMetGlnThrGlnLysGlnAlaAlaIleAsnAs 285

948 CGCCATCAACTACCTGGGGGCCGAGTCCCTGCCGCCGGTGGTGCAGACGC 997
285 nAlaIleSerTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrP 302

998 CCCCGGGCGGTTCCGAGGTGGTCCCCGTCATCAGCCCGATGTACCAGCTG 1047
302 roProGlyCysSerAsp........ValIleSerSerMetTyrGlnLeu 315

1048 CACAGG...CGCTCGGAGGGCACCCCGGCGTCCAACCACTCGGCCCAGGA 1094
316 GlnLysProGlnSerAspArgArgGlyLnProArgAsnAsnHisMetProHisAs 332

1095 CAGCGCCCGTGGAGTACCTCTGCTCGTCTCCAAGGCCAAGTTGGTGCCCT 1144
332 pGlyAlaGlyGluAsnLeuLeuLeuLeuLeuSerLysAlaLysLeuValSerS 349

1145 CGGAGGCCGCAGGCGTCCCCGACCAACAGCTGCCAAGAACTCCACGGACACC 1194
349 erGluArgAspProSerProSerAsnSerLeuGlnAspSerThrAspThr 365

1195 GAGAGCAACAACGAGGAGGACGAGGCGCCAGCCGGTCTTATCTACCTGACCAACCA 1244
366 GluSerAsnAsnGluGlu...ArgGlyGlyLeuIleTyrLeuThrAsnHi 381


1245 CATCGCCCAGCACGGCGCAACGCGTGTCGCTCCTCAAGAGGAGCACCGCGCCT 1294
     ||||||   :::   :::::  ||| |
381 sIleAlaProHisLeuArgAsnGlyLeuSerAlaLysAspLeuArgProT 398

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGAACGGCCTCCGCGTG 1344
     |||||||::::||||||||::|||:  ::::::||
398 yrAspMetLeuArgAlaGlyLyAlaAspAsnSerGlnGluProPheLysVal 414

1345 GTCAGCACCAGCGGGAGCAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     |||||::||||||||   :::::|||||||||  ||||
415 ValSerSerSerGlyLyProLeuArgMetTyrLysCysGluProCysAr 431

1395 GGTGCTCTTCCTGATCACGTCATGTCACCATCCACATGGCTGCCACG 1444
     ||||||||||||:|   ||||||||||||||||  |||
431 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 448

1445 GCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAGCCAGGAC 1494
     |||||::|||||::||||||||||||||||||||||  |||||
448 lyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 464

1495 CGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCACAT 1544
     ||||||||||||||||||||   :||| ||||||||||||||||||
465 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArgPheHisMe 481

1545 G 1545

481 t 481
```

seq_name: sp_vertebrate:Q9I8W1

```
seq_documentation_block:
ID   Q9I8W1      PRELIMINARY;       PRT;     500 AA.
AC   Q9I8W1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   IKAROS.
OS   Raja eglanteria.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes:
OC   Elasmobranchii, Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=33514;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20318730; PubMed=10861066;
RA   Haire R.N., Miracle A.L., Rast J.P., Litman G.W.;
RT   "Members of the ikaros gene family are present in early representative
RT   vertebrates.";
RL   J. Immunol. 165:306-312(2000).
DR   EMBL; AF163848; AAF87271.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00355; ZnF_C2H2; 5.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   500 AA;   55561 MW;  4DF04E1CE5BBBE84 CRC64;
```

```
alignment_scores:
           Quality:   1810.00                   Length:    530
             Ratio:      4.132                   Gaps:       10
Percent Similarity:   82.642     Percent Identity:  65.849

alignment_block:
US-08-711-417c-165 x Q9I8W1

Align seg 1/1  to: Q9I8W1  from: 1  to: 500

  1 ATGGATGCTGACGAGGTCAAGACATGTCTTTCTCATCAGG........ 42
    ||:::::||||::|||||||::||||||||:|||||||||||::|||   ||||||
  1 MetGluThrGluGluAlaGlnAspIleSerGlnLeuSerGlyGluThrGl 17
```

```
43    ....AAGGAAAGCCCCCTGTGAAGCGATACTCCAGATGAGGGCGATGAGC    88
      ::  ||||||::::       |||    |||::::|          ||
17    ySerLysGluAlaLeuSerIleIleAspGlnProGluGln...LeuGluG    33

89    CCATGCCGATCCCGGAGGACCTTCCACCACTTCGGGAGGACAGCAAAGC    138
      ||||||||||||||||||||||||||  || :::||||||||||  :::
33    lnMetProValProGluAspLeuSerThrSerAlaGlyGlyLysIleAsn    49

139   TCCAAGAGTGACAGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAG    188
      :::||| ||||:::||||||||||||||||||:::::||||||||||||
50    LeuGlnGlyAspLysValIleAlaAsnSerIleLysValGluThrProSe    66

189   TGATGAAGAGAATTGGGCTGCCTGTGAAATGAATGGGAAGAATGTGCGG    238
      ||||||||||||||||||||||||:::  ||||:::||||||||||||
66    rAspGluGluAsnGlyAlaAlaCysAlaAspThrAsnGlyLysAspCysAlaG    83

239   AGGATTTACGAATGCTTGATGCCTGGGAGAGAAAATGAATGGCTCCCAC    288
      |||||||||||||||:::  |||||||||||||||  ||||| |||||||
83    luAspLeuArgMetIleAspPheThrGlyThrLysLeuAsnGlyProHis    99

289   AGGGACCAAGGCAGCTCCGGTTTGTCGGGAGTTGGAGGCATTCGACTTCC    338
      :::   ||||||||:::||::||||||||||||||||||||||||||
100   AsnGlyGlnGlyTyrAsnLysAlaPheSerGlyValGlyValGlyIleArgLeuPr    116

339   TAACGGAAAACTAAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCA    388
      |||
116   oAsn..........................................    117

389   ATGTGCTCATGTGTTCACAAAGAAGCCACACTGGAGAACGGCCCCTTCCAG    438
      ||||||||||||||||||||||||||||||||||||||||||
118   ................................GlyGluArgProPheGln    123

439   TGCAAATCAGTGCGGGGCCTCATTCACCCAGAGGGCAACCTGCTCCGGCA    488
      |||||||||||||||||||||||||||||||||||||||||||||||
124   CysSerGlnCysGlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHi    140

489   CATCAAGCTGCATTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACT    538
      :::   ||||||||||||||||||||||||||||||||||||||||||||
140   sIleLysLeuHisSerGlyGluLysProPheLysCysHisLeuCysAsnT    157

539   ACGCCTGCCGCGAGGAGGACGCCCTCACTGGCCACCTGAGGACGCACTCC    588
      ||||||||||||||||::::||||||||||||||||||||||||||||||
157   yrAlaCysArgArgArgAspAlaLeuSerGlyHisLeuArgThrHisSer    173

589   GTTGGTAAACCTCACAAATGTGATATTGTGCCGAAGCTATAAACACGCG    638
      ||||||||||||||||||:::||||||||||||||||||||||||||
174   ValGlyLysProHisLysCysAspIleValProLysLeuAsnTyrLysSerTyrLysGlnAr    190

639   AACGTCTTTAGAGGAACACATAAAGAGCGCTGCCACAACTACTTGGAAAGCA    688
      |:::||||||||||||||||||||||||||||||||||||||::::|||
190   gSerSerLeuGluGluHisGluHisLysLysGluAlaCysHisAsnTyrLeuGlnThrM    207

689   TGGGCCTTCCGGGCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCAC    738
      |||||||| ||  :::||| ||||||||  ||||||||||||:::|||
207   etGlyLeuHisGlyGlyIleTyrSerVal...LysGluSerLysSer    222

739   AGTGAAATGGCAGAAGACCTGTTGCAAGATAGGATCAGAGAGATCTCTCGT    788
      |||||||||||||||||||||||||||||||||||:::  |:::||||||
223   HisGluLeuAlaGluGluThrGlyLysAlaSerLysLysSerGluArgAlaMetVa    239

789   GCTTGGACAGACTAGCAAGTAATGTCCGCCAAACGTAAGAGCTCTATGCCTC    838
      ||||||||||||||||||||||||||||||||||||||||||||||||||
239   lLeuAspArgLeuAlaSerAsnValAlaLaLysArgLysSerSerMetProG    256

839   AGAAATTTCTTGGGGACAAG.........GGCCTGTCCGACACGCCC    876
      ||||:::|||::::|||          ||:::||||||
256   lnLysPheValGlyGluLysThrValGlnLeuGlyIleLeuSerGluLeuGln    272

877   TACGAC...AGTGCCACGTACGAGAAGGAGGAGAACGAAAATGAAGTCCCA    923
      ||||||      |||||:::||||||||||||:::|||||:::||||||:||
273   CysAspAlaAsnSerAsnTyrGluLysGluLysGluSerGlnIleMetGlnThrHi    289
```

```
924   CGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGGCGCCGAGT    973
      :::  ::::||||||||||||||||||||||||||||||||||||||||||
289   sValLeuAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGlyAlaGluS    306

974   CCCTGCGCGCCCGCTGTGCAGACACGGCCCCGGGCGGTTCCGAGTGGTCCCG    1023
      ||||||||:::||||||||||||||||||||||||||:::||||:::|||
306   erLeuArgProLeuValGlnThrProAlaGlySerGluValIlePro    322

1024  GTCATCAGCCCGATGTACCAGCTG.........CACAGGCGGCTCGGAGGG    1064
      ||||:::|| |||||        ||          :::  ||||| ||
323   ValLeuSerAsnMetTyrProMetGlnIleThrHisSerAspThrHisGl    339

1065  CACCCCGGCGCTCCAACCACTCGGCCCAGGACCAGCGCCCTGGAGTACCTGC    1114
      ||||||:::::|| |||||||||||||||||||||||||||  ||  ||
339   yHisProAsnCys...AlaThrIleGlnAsnGlyAlaValGluAsnLeuM    355

1115  TGCTGCTCTCCAAGGCCAAGTTGGTGCCCTGGAGCCGCCAGGCGCTCCCCG    1164
      ::||||:::||||||||||     :::||||||||||||||||||||||
355   etLeuIleSerLysAlaLysSerGlnThrSerGluArgGluProSerPro    371

1165  AGCAACAGCTGCCAAGACTCCACGGACACGGAGACAACAACGAGGAGACA    1214
      |||||||||       ||||||||||||||||||||||||||||||||||
372   SerAsnSerCysProAspSerThrAspThrGluSerAsnAsnGluGlu..    387

1215  GCGCAGCCGGTCTTATCTACCTGACCAACCACCATCGCCCGACGGCGCCAAC    1264
      ||||||||||||||   |||||||||||||||||:::|||  :::||||||
388   .ArgAsnAsnLeuIleTyrLeuThrAsnHisValAlaProGlnAlaArgA    404

1265  GC.......GTGTCGTCCAAGGAGGAGGACGCACCCGCCTACGACCTGCTGCGC    1308
      ::                  |||||||||||||||||:::||||||||||
404   snGlyLeuProLeuIleLysGluGlnArgSerTyrAspLeuMetArg    420

1309  GCCGGCTTCCGAGAACTCGCAGGACGCGCTCCGCGTGTCAGCACCAGCGG    1358
      |||||  :::|| :::||||||  |||||||||||  ::::::|||:||
421   SerGlyMetGluProAsnGlnGluIleLeuIleLeuLysValValSerAlaAlaAspGl    437

1359  GGAGCAGATGAAGGTGTACAAGTGCCAAACACTGCCGGTGCTCCTTCCTGG    1408
      ||||||||||||||||||||||||||||||||||||||||||||||||||
437   yGluHisMetLysValTyrLysCysGlnHisCysArgValLeuPheLeuA    454

1409  ATCACGTCATGTAACACCACATCCCACATGGCTGCCACGGCCTTCCGTGATCCT    1458
      ||||||:::||||||||||||||||||||||||||||||||||||||||||
454   spHisValMetPheThrIleHisMetGlyCysHisLeuGlyPheArgAspPro    470

1459  TTTGAGTGCAACATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCTC    1508
      ||||||||||||||||||||||||||||||||||||||||:::||||:::|
471   PheGluCysAsnMetCysGlyTyrArgSerGlnAspArgTyrGluPheSe    487

1509  GTCGCACATAACGCCAGGGGAGCACCGCTTCCACATGAGC    1548
      ||||||||||||||||||||||||::::|||||||||||
487   rSerHisIleThrArgGlyGluHisCysTyrHisMetAsn    500
```

seq_name: sp_vertebrate:Q90XA3

Wed Aug 28 10:05:27 2002

DR   EMBL; AF416370; AAL11906.1; -
SQ   SEQUENCE 526 AA; 57510 MW; 0EBA4C49AAAF623E CRC64;

alignment_scores:
       Quality: 1762.50   Length: 535
       Ratio: 3.874   Gaps: 12
Percent Similarity: 85.047   Percent Identity: 63.738

alignment_block 1/1
US-08-711-417C-165 x Q90XA3

Align seg 1/1 to: Q90XA3   from: 1 to: 526

   1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGAAGGAAAG 50
     |||||||||||||||::::::|||::::|||||| ::|||||::::::|
   1 MetGluThrGluAlaGlnGluMetSerGlnIleThrGlyArgAspSe 17

  51 CCCC...CCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGA 97
     ::::: |||||||||::::: |||||| |||||
  17 rProMetAsnAlaAsnGluGlyGlyGluAspGlnAspGluAlaMetProV 34

  98 TCCCCGAGGACCTCTCCACCACCTCGGGAGGACGAGCAAAGCTCCAAGAGT 147
     :::||||||||||||:::::::|||||| |||::||||||::::::
  34 alProGluAspLeuSerAlaSerThrGlyLeuGlnHisAsnAsnArgThr 50

 148 GACAGAGTGCTGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGA 197
     |||: :::|| |||||||||||||||:::|||||
  51 AspLysProLeuAlaCysAsnIleLysValGluAlaArgSerAspGluGI 67

 198 GAATGGGCGTGCCTGTGAAATGAATGGG......GAAGAATGT...GCGG 238
     ||||| :::|||||||||||||| ||||
  67 uAsnGlyLeuSerCysGluMetAsnGlyGluAlaGluGluCysAlaAlaG 84

 239 AGGATTTACGAAATGCTTGATGCCTCGGGAGGAGAAAATGAATGGCTCCCAC 288
     ||||||||||||||:::::||||| ||||:||||
  84 luAspLeuArgIleLeuAspGlySerGlyAlaLysValAlaAsnGlySerHis 100

 289 AGG......GACCAAGCAGCTCGGCTTCGCGGAGTTGGAGGCATTCG 332
     ||::: :::||||||||||||||||||||
 101 AlaGlyProAspSerLysProAlaAlaTyrProThrAlaGlyGlyIleAr 117

 333 ACTTCCTAACGGAAAACTAAAGTGTGATATCTGTGGGATCATTTGCATCG 382
     ||||||||||||||||||||||||||||:::||||||
 117 gLeuLeuProAsnGlyLysLeuLysCysAspIleCysGlyIleIleCysIleG 134

 383 GGCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACTGGAGAACGGCCC 432
     |||||||||||||||||:::||||||:::||||
 134 lyProAsnValLeuMetValHisLysArgSerHisThrGlyGluArgPro 150

 433 TTCCAGTGCAATCAGTGCGGGCCTCATTCACCCAGAAGGGCAAACCTGCT 482
     ||||||||||||||||||||||||||||||||
 151 PheGlnCysAsnGlnCysGlyAlaSerPheThrGlnLysGlyAsnLeuLeu 167

 483 CCGGCACATCAAGCTGCATTCCGGGGAGAAGCCCTTCAAATGCCACCTCT 532
     |||||||||||||::|||||||||||||||||||||
 167 uArgHisIleLysLeuHisSerGlyGluLysProPheLysCysHisLeuC 184

 533 GCAACTACGCCTGCGGCCGGAGGGACGCCCTCACTGGCCACCTGAGGACG 582
     ||||||||||||||||||||||||||||||||
 184 ysAsnTyrAlaCysArgArgArgAspAlaLeuThrGlyHisLeuArgThr 200

 583 CACTCCGTTGGTAAACCTCACAAATGTGATATTGTGCCGAAGCTATAA 632
     ||||||||||||||
 201 HisSerValGlyLysProHisLysCysAspIleValProLysArgTyrLy 217

 633 ACAGCGGAAGTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGG 682
     ||||||:::||||||||||||||||||||||||
 217 sGlnArgSerLeuGluGluHisLysGluArgCysHisAsnTyrLeuG 234

 683 AAAGCATGGGCCTTCCGGCACACTGTACCCAGTCATTAAAGAAGAAACT 732
     :: ||||||::::|| :::|||||||||:::
 234 lnCysMetGlyLeuGlnAsnSerIleTyr...ThrValLysGluGluAsn 249

 733 AAGCCACAGTGAAATGGCAGAAGAGACTTGTGCAAGATGAGGATCAGAGAGATC 782
     ::::::||||::: |||||::::::||
 250 SerGlnAsnGluGlnArgGluAspMet......ProAlaSerGluArgAI 264

 783 TCTCGTGCTGGACAGACTAGCAAGTAATGTCGCAAGTAATGCCAAACGTAAGAGCTCTA 832
     |||||||||||:::||||||||::::||||||||||||||
 264 aLeuValLeuAspArgIleAlaAsnValAlaLysArgLysSerSerM 281

 833 TGGCTCAGAAATTTCTTGGGACAAGGGCCTCTCCGACACGCCCTACGAC 882
     |||||||:::|||| ||||||: :::::::
 281 etProGlnArgPheValGlyLeuAsnArgLeuSerGluLeuSerPheGlu 297

 883 AGTGCCACGTACGAGAAGGAGAACGAAATGATGAAGTCCCACGTGATGGA 932
     ||:::::::| :::||||||::||
 298 SerGlySer..........GlyGluLeuMetGlnProHisValIleAs 310

 933 CCAAGCCATCAACAACGCCATCAACTACCTGGGGCCGAGTCCTGCGCC 982
     |||||||||:::||||:::||||||||||||||||
 310 pGlnAlaIleAsnSerAlaIleSerTyrLeuGlyAlaGluSerLeuArgP 327

 983 CGCTGGTGCAGACGCCCCCGGCGCGGTTCCGAGTGGTCCCGGTCATCAGC 1032
     |||||||||::::: |||||||||:::::::||||||
 327 roLeuValGlnThrProGlySerAlaAspMet.......ValValSer 341

1033 CCGATGTACCAGCTGCACAGGCGCTCGGAGGGCACCCGCGGCTCCACCA 1082
     :::|||||||||||||||| :::::::
 342 ProLeuTyrAsnLeuHis...LysSerGlnThrAlaGluHisGlyAsnGlyVa 357

1083 CTCGGCCCAGGACAGCGCCGTGGAGTACCTGCTGCTCCAAGGCCA 1132
     |::||||||||||||:::::: ||||||||||||:::|
 357 lSerAlaLysAspSerAlaAlaGluHisLeuLeuLeuLeuSerLysSerL 374

1133 AGTTGGTGCCCTCGGAGCGCGAGGCGTCCCGAGCAACAGCTGCCAAGAC 1182
     |||||||||:::|| ||||||:::::::
 374 ysSerAlaSerValAspLysAspGlySerProSerProSerGlyGlnAsp 390

1183 TCCACGGACACCGAGCAACAACGAGGAGCAGCGCAGC........... 1221
     |||||||||||| ::: ::
 391 SerThrAspThrGluSerAsnAsnGluLysArgSerAlaGlyValSerGl 407

1222 ...........GGTCTTATCTACCTGACCAACCACCATCGCCCGAC 1255
     :::::||||||||||||||||||||||::::||
 407 yThrAlaAlaThrGlyGlyLeuIleTyrLeuThrAsnHisMetAlaProG 424

1256 GCGCGCAACGCCGTGTCCGTC..........AAGGAGGAGCACCGCGCC 1293
     :::||||||:::::::|| |||||||||||||
 424 lyMetArgAsnGlyValProGlyValValLysGluGlnGlnArgHis 440

1294 TACGACCTGCTGCGCGGCGCC..........TCCGAGAACTCGCAGGA 1331
     ::::::: ||||||||||||:::||
 441 PheGluAlaLeuArgAlaAlaGlyMetAspLeuSerIleAlaSerSerGl 457

1332 CGCGGCTCCGCGTGGTCAGCACCGGGAGCAGATGAAGGTGTACAAGT 1381
     |||:::::||||||||||||||||||||||||||||
 457 uGlyPheLysValLeuSerGlyAspGlyGluGluLeuArgAlaTyrArgC 474

1382 GCGAACACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACCATCCAC 1431
     || |||||||||||||||||||:::::||
 474 ysIleHisCysArgValLeuProHisLeuHisValMetTyrThrHisHis 490

1432 ATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTCGCGCTA 1481
     ||||||||||:::||||||||||||||||||||
 491 MetGlyCysHisGlyPheArgAspProPheGluCysAsnLeuCysGlyTy 507

1482 CCACAGCCAGGACCGGTACGAGTTCTCGTCGCCACATAACGCGAGGGGAGC 1531
     :::|||||||||||||||||||||||:::|
 507 rArgSerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluH 524

1532 ACCGC 1536
     |||||
 524 isArg 525

```
seq_name: sp_vertebrate:O93581

seq_documentation_block:
ID   O93581          PRELIMINARY;       PRT;    537 AA.
AC   O93581;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   IKAROS.
GN   IKAROS.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SPLEEN;
RA   Amemiya C., Kawasaki H.;
RT   "Characterization of zebrafish ikaros, a gene necessary for
RT   differentiation of the immune system.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AF092175; AAC61763.1; -.
DR   ZFIN; ZDB-GENE-980526-304; ikaros.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   PRINTS; PR00048; ZINCFINGER.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   537 AA;  58865 MW;  3BF047938CF38A37 CRC64;

alignment_scores:
       Quality:  1747.00          Length:    546
       Ratio:       3.840           Gaps:      13
Percent Similarity:  83.333   Percent Identity:  62.454

alignment_block:
US-08-711-417c-165 x O93581

Align seg 1/1  to: O93581  from: 1  to: 537
```

```
    1 ATGGATGCTGACGAGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG   50
      ||||||||||||||:|||:||||||       :::||||:::::||
    1 MetGluThrGluAlaGlnGluMetSerGlnIleThrGlyArgAspSe   17

   51 CCCC...CCTGTAAGCGATACTCCAGATGAGGGGATGAGCCCATGCCGA   97
      ||||        :::::: |  ::::: |||||   ||||:
   17 rProMetAsnAlaAsnGluGlyGlyGluGlnAspGlnAspGluMetProV   34

   98 TCCCCGAGGACCTCCTCCACCACCTCGGGAGGACGAGGACAAAGCTCCAAGAGT  147
      :||||||||||||||||:::::||||      |||::||||:::::::::::||
   34 alProGluAspLeuSerAlaSerThrGlyLeuGlnHisAsnAsnArgThr   50

  148 GACAGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGA  197
      ::::|| |||||||||||||||||||||||||||||||||||||||||||
   51 AspLysProLeuAlaCysAsnIleLysValGlnValGluAlaArgSerAspGluG1   67

  198 GAATGGGCGTGCCTGTGAAATGAATGGG.......GAAGAATGT...GCGG  238
      ||||||      :::||||||||| ||||        ||||||       ||||
   67 uAsnGlyLeuSerCysGluMetAsnGlyGluAlaGluGluCysAlaAlaG   84

  239 AGGATTTACGAATGCTTGATGCTCCGGGAGGAGAAAATGAATGGCTCCCAC  288
      ||||||||||||||||||:::||||||| |||||||||||:||||||||||
   84 luAspLeuArgIleLeuAspGlySerGlyAlaLysValAsnGlySerHis  100

  289 AGG......GACCAAGGCAGCTCGCTTTGTCGGGAGTTGGAGGCATTCG  332
      ||||:       :::||| ::::||||:||||||||||||||||||||||
  391 LysAspGlySerProSerGlyGlnAspSerThrAspThrGluSe  407
```

```
  101 AlaGlyProAspSerLysProAlaAlaTyrProThrAlaGlyGlyIleAr  117
      ||||||||||||||||||||||||||||||||||||||||||||||||
  333 ACTTCCTAACGACGAAAACTAAAGTGTATATCTGTGGGATCATTTGCATCG  382
      |||||||||||||||||||||||||||||||||||||||||:||||||
  117 gLeuProAsnGlyLysLeuLysCysAspIleCysGlyIleValCysIleG  134

  383 GGCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACT............  420
      |||||||||||||||||||||||||||||||||||||
  134 lyProAsnValLeuMetValHisLysArgSerHisThrGluArgLys  150

  421           GGAGAACGGCCCTTCCAGTGCAATCAGTG  449
      ||||||||||||||||||||||||||||
  151 SerValLeuGlnGlnLysGlyGluArgProPheGlnCysAsnGlnCy  167

  450 CGGGGCCTCATTCACCCAGAAGGCAACCTGCTCCGGCACATCAAGCTGC  499
      ||||||||||||||||||||||||||||||||||||||||||||||||
  167 sGlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuH  184

  500 ATTCCGGGAGGAGCCCTTCAAATGCCACCTGCAACTAGCGCCTGCCGGC  549
      ||||||||||||||||||||||||||||||||||||||||||||||||
  184 isSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArg  200

  550 CGGAGGGACGGCCCTCACTGGCCACCTGAGGACGACCACTCCGTTGGTAAACC  599
      ||||||||||||||||||||||||||||||||||||||||||||||||
  201 ArgArgAspAlaAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPr  217

  600 TCACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCAACGTCTTTAG  649
      |||||||::|||||:|||||||||||||||||||||||||::|||||||
  217 oHisLysCysAlaTyrCysGlyValArgSerTyrLysGlnArgSerLeuG  234

  650 AGGAACATAAAGAGCCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCG  699
      |||||||||||||||||||||||||||||||||||||| ||||||||
  234 luGluHisLysGluArgCysHisAsnTyrLeuGlnCysMetGlyLeuGln  250

  700 GGCACACTGTACCCAGTCATTAAAGAAGAAAACTAAGCACAGTGAAATGGC  749
      ::::::::::|| ::::||||||||  |||||||||||:::|||
  251 AsnSerIleTyr...ThrValLysGluGluAsnSerGlnAsnGluGlnAr  266

  750 AGAAGACCTGTGCAAGATAGGATCAGAGAGATCTCGTCGTGCTGGACAGAC  799
      ||||||        :::|||||||||| ::::||||||||||||||||
  266 gGluAspMet......ProAlaSerGluArgAlaLeuValLeuAspArgI  281

  800 TAGCAAGTAATGTCGCAAACGTAGAGCTCTATGCCTCAGAAATTTCTT  849
      :::||||||||||::||||||||||||||||||||::|||:||:::||
  281 leAlaAsnAsnValAlaLysArgLysSerSerMetProGlnArgPheVal  297

  850 GGGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAA  899
      ||||:::||    :::::||:::|||||||||||||||:::||::||
  298 GlyGluAsnArgLeuSerGluLeuSerPheGluSerGlySer.........  311

  900 GGAGAACGAAATGATGAAGTCCCACGTGATGACCAAGCCATCAACACG  949
      :::||||||||||||||||||::||||||||||||||||||||||::|:|
  312 ....GlyGluLeuMetGlnProHisValIleAspGlnAlaIleAsnSerA  327

  950 CCATCAACTACCTGGGGCCGACTCCCTGCGCCCGTGGTCCAGACGCCC  999
      ||||||||||||||||||||||||||||||||||||||||||:||:||
  327 laIleSerTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrSer  343

 1000 CCGGGCGGTTCCGAGGTGGTCCCGTCATCAGCCCGATGTACCAGCTGCA  1049
      ||||||      ::::||||||||||||||||||||:::||||
  344 ProGlySerAlaAspMet......ValValSerProLeuTyrAsnLeuHi  358

 1050 CAGGGCGTCCGGAGGGCACCCGGCTCCAACCACTCGGCCCAGGACAGCG  1099
      :: ::::||||    :::||     :::::| |||||||||||||
  358 s...LysSerGlnThrAlaGlyLuGlyAsnGlyValSerAlaLysAspSerA  374

 1100 CCGTGGAGTACCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAG  1149
      ||:::|||||||||||||||||||||||:::|||:||||||||||||
  374 laAlaGluHisLeuLeuLeuLeuSerLysSerAlaSerValAsp  390

 1150 CGCGAGGCGTCCCGACAACAGCTGCCACGACTCCACGGACACCGAGAG  1199
      ||::::||||||||||||||||||||||||||||||||||||
  391 LysAspGlySerProSerGlyGlnAspSerThrAspThrGluSe  407
```

Wed Aug 28 10:05:27 2002

us-08-711-417c-165.rspt

Align seg 1/1 to: Q918V9   from: 1 to: 513

```
1200 CAACAACGAGGAGGAGCAGCCCAGC...................................G 1222
     ||||||||||||||||::: ::
 407 rAsnAsnGluGluArgSerAlaGlyValSerGlyThrAlaAlaThrGlyG 424

   1 ATGGATGCTGACGAGGTCAAGACAATGTCT.......................TT 32
     ||| ::::::::::::||||:: |||                         ::
   1 MetValThrGluAspAsnGlnGlnIleSerHisProGluAlaSerPheLe 17

  33 CTCATCAGGGAAGGAAAGCCCCCTGTAAGCGATACTCCAGATGAGGGCG 82
     |||||||||||||||||||::::::: ::||::: ::|||
  17 uSerAspAlaGluThrValProThrValThrAspAlaProAsp........ 31

1223 GTCTTATCCTACCTGACCAACCACCATCGCCCGACGCGCGCAACGCGTGTCG 1272
     ||||||||||||||||||||::::::||| ::::: :: |||||
 424 lyLeuIleTyrLeuThrAsnHisMetAlaProGlyMetArgAsnGlyGly 440

1273 CTC..........AAGGAGGAGCACCGCGCTACGACCTGCTGCGCGC 1310
     |||||         ::|||:::::||| :::::: |||||||
 441 LeuProGlyValLysGluGlnGlnArgHisPheGluLeuArgAl 457

  83 ATGAGCCCATGCCGATCCCGAGGACCTCTCCACCACCTCGGGAGGACAG 132
     ::||||||                 ::::::::::::|||||||||||
  32 ..........LeuProGluLysIleAsnValAlaAspAlaValAlaValGln 43

1311 CGCC..............TCCGAGAACTGCCAGGACGACGGCCTCCGCGTGGTCA 1348
     ||||               ||| |||||||:::::::::||||||||
 457 aAlaGlyMetAspLeuSerIleAlaSerSerGlyLeuGlyPheLysValLeuS 474

 133 CAAAGCTCCAAGAGTGACAGAGTCGTGGCCAGTAATGTT................. 171
     ::::         ::::::|||||||||||||||||||             :::
  44 GluProIleAspGluLysGlnLeuSerThrAsnGlyValThrAlaGlyVal 60

1349 GCACCAGCGCGGGAGCCAGATGTACAAGTGCGAACACTGCCGGGTG 1398
     ||| ::::|||||||||||||||||  ||||||||||
 474 erGlyAspSerHisIleThrArgGlyLeuHisArg 490

 172 .........AAAGTAGAGACTCAGAGTGAT......................GAAG 196
     :::         :::||::|                              :::
  60 aValArgValTyrAspGluSerGlnIleAlaAlaLeuAspLeuSerArgSpG 77

1399 CTCTTCCTGGATCACGTCAGTACACCATCCACATGGCTGCCACGGCTT 1448
     |||||||||||||||||||||||||:::||| ||||::: :::::
 474 erGlyAspSerHisIleThrArgGlyLeuHisArg 490

 197 AGAATGGGCGTGCCTGTGAAATGAATGGGGAAGAAATGTGCGGAGGATTA 246
     ||:::|||  ::||| ||||  ::::::||||||||||||
  77 luGlyCysHisAlaGlyValGluTyrGlyGlyThrAsnSerSerCysLeu 93

1449 CCGTGATCCTTTTGAGTGCAACATGTGCGCTACCACCAGCCAGGACCCGT 1498
     ||||||||||||:::|||||||||||||||||||||||||||||||
 507 eArgAspProPheGluCysAsnLeuCysGlyTyrArgSerGlnAspArgT 524

 247 CGAATGCTTGATGCCTCGGGAGGAGAAAATGAATGGCTCCCAC...AGGGA 293
     :::|||:::                ::: :::::|||||
  94 .........ValGluAsnGluAspTyrAsnGlyAspHisPheSerAlaS 106

1499 ACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGC 1536
     |||||||||||||||||||||||||||||||
 524 yrGluPheSerHisIleThrArgGlyHisArg 536

 294 CCAAGGCAGCTTGTCGGCTTGTCTGGGATCATTTGCATCGGGCCCAATGTG 393
     ::|| |||:::|                  :::|||||||||||
 106 nGlnProSerThrAla.............AspGlyIleArgAlaSerAsnG 119

 344 GAAAACTAAAGTGTGATATCTGTGGGATCATTTCTGGATCATTTCGGGAAGT 393
     ||||||| ||||||||||| ::    ||:::|||||||||
 119 lyLysLeuThrCysAspIleCysGlyLeuSerCysValGlyProAsnVal 135

 394 CTCATGGTTCACAAAAGAAAGCCACACTGGAGAACGGCCCTTCCAGTGCAA 443
     ||||||||||||||||||||||||||||||:::|||||
 136 LeuMetValHisLysArgSerHisThrGlyGluArgProPheHisCysAs 152

 444 TCAGTGCGGGGCCTCATTCACCCAGAGAGGGCAACCTGCTCCCGGCACATCA 493
     ||||||||||||||||::::::||||||||||||||||:::
 152 nGlnCysGlyAlaSerPheThrGlnArgGlyAsnLeuArgHisValL 169

 494 AGTCGCATTCCGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCC 543
     ||||||||||||||||||||||||||||||||||||||||:::::
 169 ySLeuHisSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAla 185

 544 TGCCCCGGAGGGACGGCCCCTCACTGGCCACCTGAGGACGACCACTCCGTGG 593
     |||||||||||||||||||||||||||||||||||||||||::::::
 186 CysArgArgArgAspAlaLeuSerGlyHisLeuArgThrHisSerIleAs 202

 594 TAAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAAACAGGGAACGT 643
     |||||:::|||::::||||||||||||||||||||   :::||
 202 pLysProTyrLysCysSerPheCysSerGlyArgSerTyrLysGlnCysSerS 219

 644 CTTTAGAGGAACATAAAGAGCCGCTGCCACAACTACTTGAAAGCATGGGC 693
     ||||||||||:::||||||  :::||||||||||||||||||||
 219 erLeuGluGluHisLysAspArgCysHisSerTyrLeuGlnLysMetGly 235

 694 CTT.......CCGGGCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCA 737
     |||         ::||                      |||||
 235 ......GluGlyLeuHisHi 246

 738 CAGTGAAATGCCAGAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCG 787
     |  :::::          ::::::           ||||||||||
 246 sArgAsnLeuUlysAlaGluMet........GlySerGluArgAlaLeuV 260

 788 TGCTGGACAGACTAGCAAGTAATGTCGCAAACGTAAGAGCTCTATGCCT 837
     ||||||||||||||||||||||||||||||||||||||||||||||
 260 aLeuAspArgLeuAlaSerAsnValAlaLysArgLysSerSerMetPro 276
```

seq_name: sp_vertebrate:Q918V9

seq_documentation_block:
ID   Q918V9   PRELIMINARY;   PRT;   513 AA.
AC   Q918V9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AIOLOS.
OS   Raja eglanteria.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=33514;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20318730; PubMed=10861066;
RA   Haire R.N., Miracle A.L., Rast J.P., Litman G.W.;
RT   "Members of the ikaros gene family are present in early representative
RT   vertebrates.";
RL   J. Immunol. 165:306-312(2000).
DR   EMBL; AF163850; AAF87273.1; -.
DR   HSSP; P15822; 1BBO.
DR   InterPro; IPR000822; znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; zinc-finger.
SQ   SEQUENCE   513 AA; 56940 MW; 9D1F5CEB25B6D110 CRC64;

alignment_scores:
        Quality: 1402.50   Length: 543
        Ratio: 3.472   Gaps: 17
Percent Similarity: 74.401   Percent Identity: 54.144

alignment_block:
US-08-711-417C-165 x Q918V9   ..

OC   Rajiformes; Rajidae; Raja.
OX   NCBI_TaxID=33514;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20318730; PubMed=10861066;
RA   Haire R.N., Miracle A.L., Rast J.P., Litman G.W.;
RT   "Members of the ikaros gene family are present in early representative
RT   vertebrates.";
RL   J. Immunol. 165:306-312(2000).
DR   EMBL; AF163847; AAF87270.1; -.
DR   HSSP; P15822; 1BBO.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   522 AA; 59360 MW; AAF305808E9EA75E CRC64;

alignment_scores:
         Quality: 1396.00   Length: 493
         Ratio: 3.598   Gaps: 8
Percent Similarity: 78.702   Percent Identity: 55.984

alignment_block:
US-08-711-417C-165 x Q9I8W2

Align seg 1/1   to: Q9I8W2   from: 1   to: 522

      112 TCCACCACCTCGGGAGGACAGCAGCAAAAGCTCCAAGAGTGACAGAGTCGTGGC 161
          :::::::|| :::|||||| ||||||| || ::|||||||
       32 ThrSerThrProAsnGlyGlnTyrSerSerLeuSerHisMetThrAsnLy 48

      162 CAGTAAGTGTAAAGTAGAGACTCAGTGATGAAGAGAATGGGCGTGCCT 211
          :::::::||:|||| :||||||||:::||:||::|||::|:||:||
       48 sAsnSerIleLysIleGluMetHisSerAspGluAspGlyLysAlaL 65

      212 GT...GAAATGAATGGGGAAGAATGTCGGAGGATTTACGAATGCTTGAT 258
          ::|||:: || ||::   ::||| :|:||
       65 euLysGlnMetAspLeuAlaArgSerLysGluGluMetSerThrMetAsp 81

      259 GCCTCGGGAGGAGAAAATGAATGGCTCCACAGGGACCCAGGCAGCTCGGC 308
          ||| ::| ::: ||||||||||| ||::|||||||
       82 GluSerValValGluSerAsnGlyLeuThrGluSerAsnGlnGlyGlnGl 98

      309 TTTGTCGGGAGTTGGAGGCATTCGACTTCTCAACGGAAAACTAAAGTGTG 358
          ::::::|| |||||||||||||||||||||||||||||
       98 uMetGlnAlaGluGlyGlyGlyIleArgLeuProAsnGlyLysLeuLysCysSA 115

      359 ATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAA 408
          |||||||||||::||||||||||||||||||||||||||||
      115 spIleCysGlyMetValCysIleGlyTyrProAsnValLeuMetValHisLys 131

      409 AGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGCCTC 458
          |||||||||||:::|:||||||||||||||||||||||||||:::||||
      132 ArgSerHisThrGlyGluArgProPheHisCysAsnGlnCysGlyAlaSe 148

      459 ATTCACCCAGAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGG 508
          |||||||||||||||||||||||||||||||||:::||||||:::||||||||
      148 rPheThrGlnLysGlyAsnLeuLeuArgHisValLysLeuHisThrGlyG 165

      509 AGAAGCCCTTCAAATGCCACCTCTGCAACTACGGCCTCCCGCCGGAGGGAC 558
          ::::::||||||:::::||||||:||||||||||||||||
      165 luLysProPheLysCysProPheCysSerTyrAlaCysSerArgArgArgAsp 181

      559 GCCCTCACTGGCCACCTGGAGACGCACTCCGTTGGTAAACCTCACAAATG 608
          |||||||||||::||||||||||||||||||||||||||||
      182 AlaLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCy 198

      609 TGGATATTGTGGCCGAAGCTATAAACACGGAACGTCTTTAGAGGAACATA 658
          |:::|||||||::||||||||||||:::||||||||||||||||
      198 sAsnTyrCysGlyArgSerTyrLysGlnSerLeuSerLeuGluGluHisL 215

      838 CAGAAAATTCTTGGGACAAG............GGCCTGTCCGACACGCC 875
          ||||||||||:::|||::::|| :::|||
      277 GlnLysPheIleGlyLyGluLysGlnPheCysLeuGlyTyrGlnAspLeuGl 293

      876 CTACGACAGTGCC...ACGTACGAGAAGGAGAACGAAATGATGAAGTCCC 922
          ||| |||||| ::|||::||||||||||||:::|||
      293 nTyrThrSerAlaProSerTyrGluLysGluGluGluMetMetGlnAlaA 310

      923 ACGTGATGGACCAAGGCCATCAACAACGCCATCAACTACCTGGGGCCCAG 972
          :: ||||||||||||||||||||||||||||||:::|||||:::||||||
      310 rgPheMetAspGlnAlaAlaIleAsnAsnAlaIleSerTyrLeuGlyAlaGlu 326

      973 TCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGC...GGTTCCGAGGTCGT 1019
          ::::::|||| :::|||||||||:::::|| ::||||:::|||
      327 SerLeuArgProLeuValGlnThrProProAlaProThrAlaGluMetVa 343

     1020 CCCGGTCATCAGCCGCGATGTACCAGCGCGCGTCGCACGGGGCCACCC 1069
          ||||||||||||:::||| :::||||||||:::||||||||:::||||||
      343 lSerValIleSerSerLeuTyrProLeuSerLeuThrArgSerAspThr. 359

     1070 CGGCGCTCCAACCACTCGGCCCAGGACAGCGCCCTGGAGTACCTGCTGCTG 1119
          |||||| ::   :::::|||||:::|||||:::|||
      360 ....SerAsnGlyHisProArgValAlaAlaThrGluAsnSerGlySer 374

     1120 CTCTCCAAGGCCAAGTTGGTGCCCTCCGGAGGCGCGAGGCGTCCCCGAGCAA 1169
          ||||| ::|||||| || ||
      375 GlnValLysAlaLysLeu........GluArgArgGlyGlySerProSerAs 388

     1170 CAGCTCCGGAAGACTCCACGGACACCAGGACAGCAACACACGAGGAGCAGCGCA 1219
          |||||||||||||||||||||||||||||||||||
      388 nSerCysGlnAspSerThrAspThrArgGluSerAsnHisGlu......... 401

     1220 GCGGTCTTATCTACCTGCAACCACACATCGCCGACGCGCGCCAACCGCGTG 1269
          ::||||:::   ::|||| ::|||||||||||||::||
      402 ..GlyGlnAlaTyrLeuGlnHisHisMetMetAlaProArgPheArgAsn 417

     1270 TCGCTC........AAGGAGGAGCACCGCGCGCTACGAACCTGCTGCGCGC 1310
          :::||| |||||||||||||||||||
      418 GlyLeuAlaAlaAlaLysGluAsnProLysProPheAspIleValLysAl 434

     1311 CGCCTCCGAGAACTCGCAGGACGCGGCCTCCGCGTGGTCAGCACCAGCGGGG 1360
          ||||::::   ::|||::||||::::||||||||||||::::||||
      434 aAlaThrGlyProGlyArgGlyMetLeuLysValMetThrAsnGluGlyG 451

     1361 AGCAGATGAAGGTGTACAAGTGCAACACTGCGCGGGTGCTCTTCCTGGAT 1410
          ||||| ||||||:::|||||||||||||||::||||||||||
      451 luGlnValArgValTyrLysCysAsnHisCysAspPheArgValLeuTyrLeuAsp 467

     1411 CACGTGCATGTACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTT 1460
          ::::||||:::::|||||||||||||||||||||||||||
      468 TyrValMetPheThrIleHisMetGlyCysHisGlyPheArgAspProPh 484

     1461 TGAGTGCAACATGTGCGGCTACCACAGCCAGGACGCGTACGAGTTCTCGT 1510
          ||||::   ::::||||||:::||||||||||||||||||
      484 eGluCysAsnValCysGlyTyrArgSerArgAspArgTyrGluPheSerS 501

     1511 CGCACATAACGCGAGGGGAGCACCGCCTTC 1539
          ||:::||||||:::|||||||||||| ::|
      501 erHisIleAlaArgGlyGluHisGlyTyr 510

seq_name: sp_vertebrate:Q9I8W2

seq_documentation_block:
ID   Q9I8W2   PRELIMINARY;   PRT;   522 AA.
AC   Q9I8W2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HELIOS.
OS   Raja eglanteria.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;

```
1514 ACATAACGCGAGGGGAGCACCGCTTCCAC 1542
     |||||::|||||||||||||||||||||
 513 isIleValArgGlyGluHisThrPheHis 522

seq_name: sp_vertebrate:Q90XA2

seq_documentation_block:
ID   Q90XA2          PRELIMINARY;        PRT;      434 AA.
AC   Q90XA2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   IKAROS.
OS   Brachydanio rerio (zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RA   Willett C.E., Kawasaki H., Lin S., Amemiya C.T., Steiner L.A.;
RT   "Ikaros expression as a marker for lymphoid progenitors during
RT   zebrafish development.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF416371; AAL11907.1; -.
SQ   SEQUENCE   434 AA;   47965 MW;   9B74FBF4EB647FC2 CRC64;

alignment_scores:
                Quality:  1373.50          Length:       530
                  Ratio:     3.643           Gaps:         10
      Percent Similarity:   71.132       Percent Identity: 51.698

alignment_block:
US-08-711-417c-165 x Q90XA2

Align seg 1/1 to: Q90XA2    from: 1   to: 434

   1 ATGGATGCTGACGAGGTCAAGACATGTCTTTCTCATCAGGAAGGAAAG 50
     |::||||||||||||||||||||||||     ::|||||||||::||:
   1 MetGluThrGluAlaGlnGluMetSerGlnIleThrGlyArgAspSe 17

  51 CCCC...CCTGTAAGCGATACTCCAGTGAGGGCGATGAGCCATGCCGA 97
     ||||         ::||||::|||||||||||||  |||||   |||||
  17 rProMetAsnAlaAsnGluGlyGlyAspGlnAlaAlaMetProV 34

  98 TCCCCGAGGACCTCTCCACCACCTCGGGAGGACAGCCAAAGCTCCAAGAGT 147
     ::|||||||||||:::||||||||||:::|||      |||||||||||||
  34 alProGluAspSerHisAspArgSerAlaSerThrGlyLeuGlnHisAsnArgThr 50

 148 GACAGAGTCGTGGCCAGTAAGTTAAAGTGAGGACTTCAGAGTGATGAAGA 197
     |||::|        ::|||:::||||||||||||||||||||||||
 151 AspLysProLeu...................................Ser 429

 198 GAATGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTAC 247
     :::||||      ::|||||||||||||||||||||
 430 ProAlaTyrThrLysGluLeuAlaLeuGluAlaThrLysValLe 446

 248 GAATGCTTGATGCCTCGGGAGAGAAATGAATGGCTCCCCACAGGGACCAA 297
     ::|||||||||        |||||||||||||||||:::||||||||
 463 uSerHisThrSerLysGlyIleTyrArgValPheSerGluGlyGluG 463

 298 GGCAGCTCCGGCTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAA 347
     |:::||||||||||||||||||||||||||||||||||
 464 lnIleArgAlaPheLysCysGluHisCysSerArgValLeuAspHis 479

 348 ACTAAAGTGTGATATCTGTGGATCATTTGCATTCGGCGGCCCAATGTGCTCA 397
     ||||||||||||||||||||||||||||||||
 480 ValMetTyrThrIleHisMetGlyCysHisSerGlyHisPheArgAspProPheGl 496

 398 TGGTTCACAAAGAGAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAG 447
     |||||||||||||||||||||||||||||||||||||||||
 496 uCysAsnIleCysGlyTyrArgSerHisAspArgTyrGluPheSerH 513
```

```
 659 AAGAGCCGTGCCACAACTACTTGGAAAGCATGGCCTTCGGGACACACTG 708
     |||||||||||||||||||||::::::  |||:  :::::::
 215 ysGluArgCysHisAsnTyrLeuGluSerMetGlyMetGlnAlaAlaGln 231

 709 TACCCAGTCATTAAAGAAGAAACTAAG...CACAGTGAAATGGCAGAAGA 755
     ::::: :: |||| ||| :: ::|||     ||||| :::
 232 ProMetAsnHisGlnAspAspTyrLysAspHisAsnThrAlaAlaAspGl 248

 756 CCGTGCAAGATAGGATCAGAGAGATCTCTGCTGTGCTGGACAGACTAGCAA 805
     ||||  |||||  :|||||:::|||: :::||||||||
 248 nIleProLeuMetGlnPheGluArgProAlaValIleGluArgLeuAlaS 265

 806 GTAATGTCCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGAC 855
     ||||||::|||||||||||||||  |||||::::|||
 265 erAsnMetGlyLysArgLysSerSerThrProGlnLysPheValGlyGlu 281

 856 AAGGGCCTGTCCGACACGCCCTACGAC.............AGTGCCAC 890
     |||  ::                      |||
 282 LysPheMetArgLeuAsnTyrProAspIleAsnCysAsnMetAsnLeuAs 298

 891 GTACGAGAGGAGAACGAAATGATGAAGTCCCACGTGATGACCAAGCCA 940
     ||||||||||||||||||::::::|||||||||
 298 nTyrGluLysGluAlaGluPheMetHisSerHisMetMetAspGlnAlaI 315

 941 TCAACAACGCCATCAACTACCTGGGCGAGTCCTGCCCGTGTG 990
     |||||||||||||||||||||||||||||||||||||||||
 315 leAsnAsnAlaIleThrTyrLeuGlyAlaGluSerLeuArgProLeuMet 331

 991 CAGACGCCCCCGGCGGT...TCCGAGTTGGTCCCGGTCATCAGCCCGAT 1037
     |||        |||:::     |||:||||||||||||:::  ::
 332 GlnHisGlnProSerMetAlaGluValValProValMetAsnSerVa 348

1038 GTAC...CAGCTGCAC.......AGGGCCTCGGAGGGCACCCCGCGCT 1075
     ||||      ::: |||          :: |||  ||||
 348 lTyrSerGlnIleTyrHisProThrArgLeuArgProSerSerArg 365

1076 CCAACCACTCGGCCCAGGACAGCGCCGTGGACTACCTGCTGCTCTCTCC 1125
     ::                  ||||:::    ||| ||
 365 luIleThrGluSerHisGluAsnThrValAspGlyProIleSerLeuIle 381

1126 AAGGCCAAGTTGGTGCCCTCGGAGCGCGAGGCGTCCCAGCAACAGCTG 1175
     ::||||||        ::|||||||||||||||
 382 ArgAlaLysAsnProThrThrGluArgValMetSerProSerHisSerCy 398

1176 CCAAGACTCCACGGACACCGAGAGCAACAACGAG............GAGC 1213
     | |||||||||||||||||||            :::|
 398 sLeuAspSerThrAspThrGluSerSerHisAspSerProArgSerAlaGlnG 415

1214 AGCCCAGCGGTCTTATCTGACCAACCACCATCGCCCGACGCCGCGCAA 1263
     ||::: |||:|     ||              :::||||
 415 lnHisProGlyValProAlaLeuThrProGlnSerLysArg.....Ser 429

1264 CGGCTGTCCGCTCAAGGAGGAGCACCGCCTGCTACGCTGCTGCCGCCGC 1313
     :::           ||||||       :: ||||||
 430 ProAlaTyrThrLysGluLeuAlaLeuGluAlaThrLysValLe 446

1314 CTCCCAGAACTCGCCAGGACGCGCTCCGCTGGTCAGCCACCAGCCGGGGAGC 1363
     ::::::                   ||  ::||||||||
 446 uSerHisThrSerLysGlyIleTyrArgValPheSerGluGlyGluG 463

1364 AGATGAAGGTGTCAACAAGTGCCAACACTGCCGGGGTGCTCTTCCTGGATCAC 1413
     ||:::||||||||||||||||||||||||
 463 lnIleArgAlaPheLysCysGluHisCysSerArgValLeuAspHis 479

1414 GTCATGTACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGA 1463
     |||||||||||||||||||||||||||||||||||||
 480 ValMetTyrThrIleHisMetGlyCysHisSerGlyHisPheArgAspProPheGl 496

1464 GTGCAACATGTGCGGCTACCACAGCCAGGAACCGGTACGAGTTCTCGTCGC 1513
     ||::||||||||:::|||||||||||||||||
 496 uCysAsnIleCysGlyTyrArgSerHisAspArgTyrGluPheSerSerH 513
```

```
  55 ...................GlyGluArgProPheGlnCysAsnGln 63

 448 TCGGGGGCCTCATTCACCCAGAAGGCAACCTGCTCCGGCACATCAAGCT 497
     ||||||||||||||||||||||||||||||||||||||||||||||
  64 CysGlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLe 80

 498 GCATTCCGGGGAGAAGCCCTTCAAATGCCACCCTGCAACTACGCCTGCC 547
     ||||||||||||||||||||||||||||||||||| |||||||||||
  80 uHisSerGlyGluLysProPheLysCysHisLeuLeuAsnTyrAlaCysA 97

 548 GCCGGAGGACGGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAA 597
     |||||||||||||||||||||||||||||||||||||||||| ||||:||
  97 rgArgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLys 113

 598 CCTCACAAATGTGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTT 647
     ||||:|||||:||::|||||||||||||||||||::||||||||:|||
 114 ProHisLysCysAlaTyrCysGlyTyrArgSerTyrLysGlyAsnSerLe 130

 648 AGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTC 697
     |||||||||| |||||||||||||||||||||||||||||||||||||||
 130 uGluGluHisLysGluArgCysHisAsnTyrLeuGlnCysMetGlyLeuG 147

 698 CGGGCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATG 747
     :::::::::||       :::::||||||||::::::::::|||:||:||:
 147 lnAsnSerIleTyr...ThrValLysGluGluAsnSerGlnAsnGlnGln 162

 748 GCAGAAGACCTGTCAAGATAGGATCAGAGAGATCTCTGTGCTGGACAG 797
     ||||||||||:|||||||:|||||||||:|||||:||||:|||||||||
 163 ArgGluAspMet......ProAlaSerGluArgAlaLeuValLeuAspAr 177

 798 ACTAGCAAGTAATGTCGCCAAACGTAAGAGTCTATGCCTCAGAAATTTC 847
     |::::||:|||||||||||||||||||||||||||||||||||||||||
 177 gIleAlaAsnAsnValAlaLysArgLysSerSerMetProGlnArgPheV 194

 848 TTGGGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAG 897
     :::::::::||||||  :::::||||||||||||||||||||||||||||
 194 alGlyAsnArgLeuSerGluLeuSerPheGluSerGlySer...... 208

 898 AAGGAGAACGAAATGATGAAGTCCCACGTGATGACCAGCCATCAACAA 947
     ::|||::::|||  ::::::||||||||||||||||||||:|||||||:|
 209 ......GlyGluLeuMetGlnProHisValIleAspGlnAlaIleAlaAsnSe 223

 948 CGCCATCAACTACCTGGGGCCGAGTCCCTGCGCCCCGTTGCGTGCGACGGC 997
     :||||||||||||||||||||||||||||:||||||||||||||||||||
 223 rAlaIleSerTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrS 240

 998 CCCCGGCGGTTCCGAGGTGTGTCCCGGTCATCAGCCCGATGTACCAGCTG 1047
     :::::::||||||||||||  :::::|||||||||||||||||||||:||
 240 erProGlySerAlaAspMet......ValValSerProLeuTyrAsnLeu 254

1048 CACAGGCGCTCGGAGGGCACCCGCGGTCCAACCACTCGGCCCAGGACAG 1097
     |||||:|||||| |||   :::::||:|||::::|||||||||||||||
 255 His...LysSerGlnThrAlaGlyLeuAsnGlyValSerAlaLysAspSe 270

1098 CGCCGTGGAGTACCTCGTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
     |||||:::|||||||||:||||:::::|||||||||||  ::::|||
 270 rAlaAlaAlaGluHisLeuLeuLeuLeuSerLeuSerSerLysSerValA 287

1148 AGCGGGAGGCGGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     :::::::::|||::::::|||||||   ||:|||||||||||||||||
 287 spLysAspGlySerProSerProGlyGlnAsnGlnLeuMetAlaGlnMetProGlyArgAspThrGlu 303

1198 AGCAACAACGAGGAGCAGCGCAGC.......... 1221
     |||||||||||||  :::
 304 SerAsnAsnGluGluGluAlaGlyArgSerAlaGlyLeuValValSerGlyThrAlaAlaThrGl 320

1222 .GGTCTTATCTACCTGACCAACCAACCACATCGCCCGCGCGCGCACGCGCGTGT 1270
     ||||||||||||  :::::  ::::::||||||||||
 320 yGlyLeuIleTyrLeuThrAsnHisMetAlaProGlyMetArgAsnGlyGlyG 337

1271 CGCTC.......AAGGAGGAGGCACCGCCGCTACGACCTGCTCGCGC 1308
     :|||                           |||||||||||
 337 lyLeuProGlyValLysGluGluGlnGlnArgHisPheGluAlaLeuArg 353
```

```
1309 GCCGCC..........TCCGAGAACTCCGAGACGGCGCTCCGCTGGT 1346
     |||||              |||    :::::::::::::::::::|:
 354 AlaAlaAlaGlyMetAspLeuSerIleAlaSerGluArgGlyIlePheLysValIle 370

1347 CAGCACCAGCGGGGAGCCAGATGAAGGTGTCAAGTGCGAACACTGCCGGG 1396
     :||| ::|||:|||||||||||||||||||||:|||| ||||||||||||
 370 uSerGlyAspGlyGluGluuArgAlaTyrArgCysIleHisCysArgV 387

1397 TGCTCTCCTGGATCACGTCATGTACACCATCCACCATGGGCTGCCACGGC 1446
     ||||||||||||||||||||||||||:|||||||||||||||||||||||
 387 aLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisGly 403

1447 TTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAGCCAGGACCG 1496
     ||||:|||||||||::|||||||::|||||:||||||||||||||||||
 404 PheArgAspProPheGluCysAsnMetCysGlyTyrArgSerGlnAsnAr 420

1497 GTACGAGTTCTGTCGCCACATAACGCGAGGGAGCACCGC 1536
     |||||||||||:|||||||:||:||||||||||||||||
 420 gTyrGluPheSerHisIleThrArgGlyGluHisArg 433
```

seq_name: sp_vertebrate:Q90WJ6

alignment_scores:
               Quality: 1351.50   Length: 522
               Ratio: 3.643   Gaps: 9
Percent Similarity: 71.073   Percent Identity: 51.724

alignment_block:
US-08-711-417C-165 x Q90WJ6

Align seg 1/1 to: Q90WJ6 from: 1 to: 428

```
   1 ATGGATGCTGACGAGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
     ||||:      ::|||||||||||||||||||||:::::::::||
   1 MetGluMetGluGluAlaGlnGluMetAlaGlnMetProGlyArgAspSe 17

  51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGGATGAGCCCATGCCGATCC 100
     ||||||::::::::::  :||||||:::::::::::::::|||||||| ::
  17 rProProAlaAsnGluAlaSerGluGluAlaGluGluProMetAlaValP 34

 101 CCGGAGGACCTCTCCACCACCTCGGGAGGACAGCAGCAAAGCTCCAAGAGTGAC 150
     ||||||||||  :::::::::  ::::::::::::::  :||||||| :::|
  34 roGluAspLeuSerAlaGlySerAlaHisSerGlnGlnAsnAsnArgGlyAsp 50

 151 AGAGTCGTGGCCAGTAATGTTAAAGTACGAGACTGATGAGAGAA 200
     :::                       ::::::::::::::::
  51 Lys............................................ 51
```

Wed Aug 28 10:05:27 2002

```
201 TGGGCGTGCCTGTGAAATGAATGGGAAGAATGTGCGGAGGATTTACGAA 250

51  .................................................. 51

251 TGCTTGATGCCTCGGGAGAGAAAAATGAATGGCTCCCACAGGGACCAAGGC 300

51  .................................................. 51

301 AGCTCGGCTTTGTCGGGAGTTGGAGGCCATTCGACTTCCTAACGCGAAAACT 350

51  .................................................. 51

351 AAAGTGTGATATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400

51  .................................................. 51

401 TTCACAAAAGAAGCCACACTGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
    |||||||||||||||::||::::||||||
52     GlyGluArgProPheHisCysSerGlnCys 61

451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTCCA 500
    |||||||||||||| ||||:::||||||||||||||||||
62  GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 78

501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGC 550
    |||||||||||||||||| |||::||||||||||||||||
78  sSerGlyGluLysProPheLysCysHisArgCysSerTyrAlaCysArgA 95

551 GGAGGACGGCCCTCACTGGCCACCTGAGGACGACGCACTCCGTTGGTAAACCT 600
    |||||||||||||||||||||||||||||||||||||||||
95  rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 111

601 CACAAATGTGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
    |||||||||::::|||| :||||||
112 HisLysCysAlaTyrCysGlyArgSerTyrGluGlnArg...SerLeuAs 127

651 GGAACATAAAAGACGGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
    :|||||||||||||||||||::: ||||||:: :
127 pGluHisLysGluArgCysHisAsnTyrLeuGlnCysMetGlyPheGlnA 144

701 GCACACTGTACCCAGTCATTAAAGAAGAAGAAACTAAGCACAGTGAAATGGCA 750
    :::||:::||| ||||:||||||:||
144 snSerIleTyrThrValValLysGluGluSerAsnGlnAsnGluGlnLysArg 160

751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
    |||||||| ::: |||||||||:||||||||||||||||
161 GluAspLeuSerGlnThrGlySerAspArgAlaLeuValLeuAspPArgLe 177

801 AGCAAGTAATGTGCCAAAGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
    |||::||||||||||||||||::||||||||||||||||||::::|
177 uAlaAsnAsnValAlaLysArgLysSerThrMetProGlnLysPheValG 194

851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
    |||||||||||||||::: ||||||:::
194 lyAspLysArgLeuSerAspLeuSerTyrAspSerGlyGlyAla....... 207

901 GAGAAGCGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
    ::||||:::|||| |||||||||||||||||:||
208 ...GlyGluLeuIleGlnProHisValIleAspGlnAlaIleAsnSerAl 223

951 CATCAACTACCTCGGGGGCCCAGTCCCTGCGCCCCGCTCGTGCAGACGGCCCC 1000
    |||::|||||||||||||||||||||||||||||::::|
223 aIleSerTyrLeuGlyAlaGluProLeuValGlnThrSerP 240

1001 CGGGCGGTTCCGAGGTGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
     ||::::||| ||| :||:: ||||||
240 roAlaSerSerThrAspValGly...LeuGlySerMetTyrProHisHis 255

1051 AGG...CGCTCGGAGGGC...ACCCCGCGCTCCAACCACTCGGCCCAGGA 1094
     ::: :::::||| ::: :|||||::||
256 LysProAlaAsnAspGlyHisValGlyArgProSerLeuSerAlaLysAs 272

1095 CAGCGCCGGTGGAGTACCTGCCTGCTGTCTCTCCAAGGCCAAGTTGGTGCCCT 1144
```

```
     ||||||::||| ||||||||||||::::::|||::::|||         :::|
272 pSerAlaAlaGluAsnLeuLeuLeuLeuSerAsnSerLysSerAlaSerS 289

1145 CGGAGGCCGAGGCGTCCCGAGCAACAGCTGCCAAGACTCCACGGACACC 1194
     ||||::::::|:||||||:::||| ||||||||||||
289 erGluLysAspGlySerProSerHisSerGlyGlnAspSerThrAspThr 305

1195 GAGAGCAACACGAGGAGCAGCGCAGCGGT.............CTTATCTA 1232
     |||||||:::: :::||| ::|||              ||||||
306 GluSerAsnAsnGluAspArgThrGlyGlyProProSerLeuIleTy 322

1233 CCTGACCAACCACATC..........GCCCGACGGCGCCAACGCGTGTCGC 1273
     ||||:::||||||:            :|||:::::::|             :
322 rLeuThrSerHisIleAsnProGlyValArgAsnGlyValLeuProLeuV 339

1274 TCAAGGAGAG...CACCGGCGCTACGACCTGCTGCGCCGCCCTCCGAG 1320
     ::||||||||    :::||| |||:: ::|||||||||
339 alLysGluGlnGlnArgGlnTyrGluAlaIleArgValIleArgMetGlu 355

1321 AACTCGCAGGACGCGCTCCGGTGGTCAGCACCAGCCGGGAGCAGATGAA 1370
     :::::::||||||||||||||:::|||::|||||||||||
356 MetAlaSerGluGlyPheLysValThrAlaAspGlyGluGlnValTh 372

1371 GGTGTACAAGTGCGAAACTGCCGGGTCTCTTCCTGATCACGTCATGT 1420
     |:::::|||:||||||||||||||||::|::|||||||||||
372 rAlaTyrArgCysGluHisCysSerArgValLeuPheLeuAspLeuValMetT 389

1421 ACACCATCCACATGGCGTGCCACGGCTTCCGTGATCCTTTTGAGTGCAAC 1470
     |||||||||||||||||||||||||||||||||||||||
389 yrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluCysAsn 405

1471 ATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCTCGCCACATAAC 1520
     :::::||||||||||||||||||||||||||||||||||||
406 LeuCysGlyHisArgSerGlnAspArgTyrGluPheSerSerHisIleTh 422

1521 GCGAGGGGAGCACCGC 1536
     ||||| ||||||
422 rArgGlyValHisArg 427
```

seq_name: sp_vertebrate:Q90W82

alignment_scores:
     Quality: 1341.50   Length: 581
     Ratio: 3.240   Gaps: 18
Percent Similarity: 71.256   Percent Identity: 49.225

alignment_block:
US-08-711-417C-165 x Q90W82

Align seg 1/1 to: Q90W82   from: 1   to: 563

```
745  ......ATGGCAGAAGACCTGCTGCAAGATAGGATCGAGAGATCTCTGT  788
          :::::::      |||  :::      |||||
286  ProValMetAspAsnAsnIleProMetValProPheGluArgProAlaVa  302

789  GCTGGACAGACTAGCAAGTAATGTCCCAAACGTAAGAGTCTATGCCTC  838
     ::|||||:      |||  |||  ||||||||||||
302  IleGluLysLeuThrSerAsnLeuGlyLysArgLysSerSerThrProG  319

839  AGAAATTCTTGGGACAAG...........GGCCTGTCCGACACGCCC  876
     |||||:  |||:||  |||                |||
319  lnLysPheValGlyGluLysLeuMetArgLeuGlyTyrProAspIleHis  335

877  TACGAC..AGTGCCACGTACGAGAGGAAGAGAACAAAATGATGAAGTCCCA  923
     :::||     :::     ::|||||:|||||:||  |||||:
336  PheAspMetAsnLeuSerTyrGluLysGluLeuGlnSerGl  352

924  CGTGATGGACCAAGCCATCAACACGCCATCAACTACCTGGGGCCGAGT  973
     ::||||||:|||||||  ||::|||:::::|||  |||||||||||:
352  nMetMetAspGlnAlaIleHisAsnAlaIleAlaTyrLeuGlyAlaAsps  369

974  CCCTGCGCCCGCTGGTGCAG...ACGCCCCGGCGGTTCCAGGTGGTC  1020
     ||||||::|||:|||||      |||||||||:||:::||||||:
369  erLeuHisProLeuMetGlnHisThrProSerThrIleAlaGluValAla  385

1021  CCGGTCATCAGC...........CCGATGTACCAGCTGCACAGGCCGCTC  1058
      |||||  |||              ::|||||:|||||||||::
386  ProValSerSerProTyrAlaGlnValTyrHisProAsnArgIleGl  402

1059  GGAGGGCACCCCGCTCCAACCACTCGGCCAGGCAGCGCCGTGGAGT  1108
      :::       |||:|||||||||||||||:||||||||||:
402  uArgProIleSerArgGluThrAlaAspSerHisGluAsnAsnMetAspG  419

1109  ACCTGCTGTCTGCTCTCCAAGGCCAAGTTGGTGCCCTGGAGCGGGAGGCG  1158
      ::|      |||  |||      ||||||||||||:||||||:|
419  lyProIleSerLeuIleArgProLysSerArgThrGlnAspArgGluGly  435

1159  TCCCCGAGCAACAGCTGCCAAGACTCCACGGACCACCGAGCAACACGA  1208
      |||||||||||||||||      |||||||||||||||||||||||
436  SerProSerAsnSerCysLeuAspSerThrAspSerGluSerHisGl  452

1209  GGAGCAGCGCAGCGGTCTTATCTACCTGACCAAC........CACATCG  1249
      :::::::::|||              |||||||||  :
452  uAspArgGlnSer.........TyrGlnGlyAsnSerAlaLeuAsnProL  466

1250  CCCGACGCGCGAACGGTGTCGCTCAAGGAGGAGGCACCGGCCTACGAC  1299
      |||::  :::   ::|||||||||||||||:|||||:  ::|||  ||
466  ysArgLysProSerProAlaTyrMetLysGluAspAlaLysAlaLeuAsp  482

1300  CTGCTGCGCGCCCGCCGAGAACTCGCAGGACGGCGCCTCCGGTGGTCAG  1349
      ::|::  :::      ||||||      ::  |||||||||||||
483  AlaThrLysAlaSerLysSerGlySerLeuLysAspIleTyrLysValIleAs  499

1350  CACCAGCGGGGAGCAGGAGCGCAGGATGAAGGTGTACAAGTGCGAACTGCCGGGTGC  1399
      :::||||||||||||||||||||||||||||||||||||:
499  nGlyGlyGluGlyGlnIleArgLeuArgAlaPheLysCysGluHisCysArgValL  516

1400  TCTTCCTGGATCACGTCATGTACACCATCCACATGGGCTGCCACGGCTTC  1449
      |||||||||||:||||||::||||||||:|:||||||||||||||
516  euPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisGlyTyr  532

1450  CGTGATCCTTTGAGTGCAACATGTGCGGCTACCACACGACCGGTA  1499
      ||||||||||:||||:||||||||||||||||||||||||||
533  ArgAspProLeuGluCysAsnIleCysGlyTyrArgSerGlnAspArgTy  549

1500  CGAGTTCTCGCGCCACATAACGCGAGGGGAGCACCGCTTCCAC  1542
      ||||||||:|:|||||||||||||||||||||  |||||
549  rGluPheSerSerHisIleValArgGlyGluHisThrPheHis  563
```

```
1  ATGGATGCTGACGAGGTCAAGACATGCTTCTCCATCAGGGAGGAAAG  50
   ||::::|||      |||      :::::  :||:::|||  |||||
1  MetGluAla...GluAlaAlaAspGlyTyrIleThrSerGlyArgAspAsnGl  16

51  CCCCCTGTAAGCAGTACTCCAGATGAGGGCGATGAGCCCATGCCGATCC  100
    ::|||||:::      ||  |||  :::||::|||||
16  uLeu.......SerProGluArgGluHisSerSerMetAlaIle.  28

101  CCGAGGACCTC...TCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGT  147
     |||||:  ||||||      |||:|||:|||:|:||
29  ....AspLeuThrSerSerThrProAsnGlyGlnHisThrSerProSer  43

148  GACAGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGA  197
     ::  ::::::::|||||  ||||||||||||||||||||||
44  HisValAlaSerThrAsnServalLysValGluMetGlnSerAspGluGl  60

198  GAATGGGCGCTGCCTGTGAAATAAGAATGGGGAAGAAATGTGCGGAGGAT...T  244
     ::|||||:||||||  :||:||  :||||:||||||  |||  :
60  uSerAspArgLys..........ThrPheCysGlyGluAspGluV  72

245  TACGAAGTGTTGATGCCTCGGAGAGAAAAATGAATGGGCCTCCCACAGGGAC  294
     ::|||::::|||||||  ::||||      ||  |||||  :::
72  alArgValLeuAsp...GluGlySerLeuGlyLeuGluProLeuThrGlu  87

295  CAAGGCAGCTCGGCT..............TTGTCGGGAGTTGG  323
     :::::::  |||                  ||  |||:::
88  AsnAsnGluMetAlaAspAsnArgLysIleGlnGluLeuSerGluGl  104

324  AGGCATTCCACTTCCTAACGGAAAACTAAAGTGTGATATCTGTGGGATCA  373
     |||||:::||||:||      :::  ||||||||::|||||::|::
104  yGlyIleArgLeuProAsnGlyLysLeuLysCysAspValCysGlyMetV  121

374  TTTGCATCGGGCCCAAGTGCTCCATGTTCACAAAAGAAGCCACACTGGA  423
     :|||||||||||||||||||||||||||||||||||||||||
121  alCysIleGlyProAsnValLeuMetValHisLysArgSerHisThrGlyI  137

424  GAACGGCGCCTTCCAGTGCAATCAGTGCGGGGCCTCATTCACCCAGAAGGG  473
     ||||||||||||||||||||||||||||||||||||||||||
138  GluArgProPheHisCysAsnGlnCysGlyAlaSerPheThrGlnLysGl  154

474  CAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGAAGCCCTTCAAAT  523
     ||||||||||||||||||||||||||||||||||||||||||
154  yAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLysProPheLysC  171

524  GCCACCTCTGCAACTACGCCGTCCGCCGGGAGGACGCCCCTCACTGGCCAC  573
     ||  :::||||||||||||||||||||||||||||||||||
171  ysProPheCysSerTyrAlaCysArgArgArgAspAlaLeuThrGlyHis  187

574  CTGGAGGACGACTCCGTT..........................591
     ||||||||||||||||:|||
188  LeuArgThrHisSerValProGlyIleGluMetAlaLeuLeuGlnGluGlnIl  204

591  ....................................591
204  eProValGlnArgLysProLeuAsnTyrSerSerGluLeuLeuTyrMetS  221

592  .................GGTAAACCTCACAAATGTGGATAT  615
     ||||||||||||::|||
221  erGlnAspLeuProSerThrMetGlyLysProHisLysCysAsnTyr  237

616  TGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCG  665
     |||||||||||||||||||||||||:||||||||||||||||
238  CysGlyArgSerTyrLysGlnArgSerSerLeuGluGluHisLysGluAr  254

666  CTGCCACAACTACTTGGAAAGCATGGGCGCCTTCCGGGCACA..........  705
     |||||||||||||||:::|:||  :::::::
254  gCysHisAsnTyrMetGlnAsnValGlyMetGluAlaGlyValGlnValL  271

706  ......CTGTACCCCAGTCATTAAAGAAGAAAACTAAGCACAGTGAA  744
     :::  |||::|||      ||:::|||||
271  euSerHisHisValProProMet......GluAspCysLysGlySerGlnGlu  285
```

seq_name: sp_human:Q9UKT9

us-08-711-417c-165.rspt

Wed Aug 28 10:05:27 2002

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ZINC FINGER DNA BINDING PROTEIN AIOLOS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LYMPHOMA;
RX   MEDLINE=20021771; PubMed=10552935;
RA   Hosokawa Y., Maeda Y., Takahashi E.i., Suzuki M., Seto M.;
RT   "Human aiolos, an ikaros-related zinc finger DNA binding protein: cDNA
RT   cloning, tissue expression pattern, and chromosomal mapping.";
RL   Genomics 61:326-329(1999).
DR   EMBL; AF129512; AAF13493.1; -.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE    509 AA;   57943 MW;  28296F1AE2F82450 CRC64;
```

```
alignment_scores:
         Quality: 1303.50                   Length:  512
           Ratio:    3.430                     Gaps:   14
Percent Similarity: 74.219     Percent Identity: 54.883

alignment_block:
US-08-711-417C-165 x Q9UKT9     ..

Align seg 1/1  to: Q9UKT9   from: 1  to: 509
```

```
22  GACATGTCTTTCTCATCAGGGAAGGAAAAGCCCCCTGTAAGGCGATACTCC  71
                               |||            |||
28  AspTyrSerLeuThrLysSerHisGluAsnVal...........          40

72  AGATGAGGGCGATGAGCCCATGCCGATCCCGAGGACCTCCACCACCT      121
    ||:||::||   |||        |||||::
41  .AspSerGlyGlyLeuProAlaAsnGluAspIle...........         53

122 CGGGAGCAGCAGAAGCTCCAAGAGTGACAGAGTCGTGGCCAGTAATGTT    171
    |||:::|||||||  |||
54  ....GlyAspAspSerMetLys...........                     60

172 AAAGTAGAGACTTCAGACTGATGAAGAAGAATGGGCGTGCCTGTGAAAATGAA 221
    |||    :::||||||        ::||||||
61  LysAspGluTyrSerGlyLeuArgAspGluAsnValLeuLysSerGluProMe 77

222 TGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATGCCTCGGGAGAGA   271
    ||:::  |||||||   :::  |||||||
77  tGlyAsn......AlaGluGluProGluIleProTyrSerTyrSerArg    92

272 AAATGAATGGCTCCCACAGGACCAAGGCAGCCAGCTCGGCTTTGTGGGGAGTT 321
    :::        ||||:::        :::
92  luTyrAsnGluTyrGluAsnIleLeuLysLeuGluArgHisValValSerPhe 108

322 GGAGGCCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTGTGGGAT  371
    :::        :::||||:||:||::|||||:
109 AspSerSerArgProThrSerGlyLysMetAsnCysAspValCysGlyLe   125

372 CATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACTG   421
    ||||||:||  |||||||
125 uSerCysIleSerPheAsnValLeuMetValHisLysArgSerHisThrG   142

422 GAGAACGCGCCTTCCAGTGCAATCAGTGCGGGCCTCATTCACCCAGAAG    471
    |||||||||||||||||||||||||||||  |||||||||||
142 lyGluArgProPheGlnCysAsnGlnCysGlyValAlaSerPheThrGlnLys 158

472 GGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGAGGAAGCCCTTCAA   521
    ::||||||  |||                        :::||||||||||||
159 GlyAsnLeuLeuArgHisIleLysLeuHisThrGlyGluLysProPheLy   175

522 ATGCCACCTCGCAACTACGCCTGCCGCGAGGACGGACCCCTCACTGGCC    571
    |||||||||||||:::|||||||||||:::||||||||||||||||
175 sCysHisLeuCysAsnTyrAlaCysArgArgAspAlaLeuThrGlyLeuH   192

572 ACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGTGATAATTGTGGC   621
    ||||||||||||||  ||  :::||||||  |||||||||||
192 isLeuArgThrHisSerValGluLysProTyrLysCysGluPheCysGly   208

622 CGAAGCTATAAACAGGGAACGTCTTTAGAGAACATAAAGAGGCGCTGCCA   671
    |||||||||||:|||||||||||:::|||||||||||||||||:::
209 ArgSerTyrLysGlnArgSerLeuGluGluHisLysGluArgCysAsr     225

672 CAACTACTTGGAAAGCATGGGCCTTCCGGCACACTGTACCCAGTCATTA    721
    ::::::|||:|||        |||||
225 gThrPheLeuGlnSerThrAsp...ProGly...............      234

722 AAGAAGAAACTAAGCACAGTGAAATGGCAGAGAGACCTGTGCAAGATAGGA  771
    ||||||||||||||||||||||||||||||||||||||||||||||
235 .....AspThrAlaSerArgAlaGluAlaArgHisIleLysAlaGluMetGly 249

772 TCAGAGATCTCTCCGTGCTGGACGACTAGCCAGTAATGTCGCCAAACG    821
    |||||::||||||||||||||||||||||||||||:::
250 SerGluArgAlaLeuValLeuAspArgLeuAlaSerAsnValAlaLysAr   266

822 TAAGAGCTCTATGCCTCAGAAATTCTTGGGACAAG...GGCCTGTCCG    868
    ||||||||||||||||||::||:::||:||
266 gLysSerSerMetProGlnLysPheIleLeuGlyGlnLysArgHisCysPheA 283

869 ACACGGCCTACGACAGTGCC...ACGTACGAGAAGGAGAACGAAATGATG   915
    ||:::  :::||:::|||
283 spValAsnTyrAsnSerSerTyrMetTyrGluLysGluSerGluLeuIle   299

916 AAGTCCCACGTGATGACCAAGCCATCAACAACCGCCATCAACTACCTGGG   965
    :::::|||||||||||||||||||||||:::|||
300 GlnThrArgMetMetAspGlnAlaIleAsnAsnAlaIleSerTyrLeuGl   316

966 GGCCGAGTCCTGCGCCCGTGGTGCAGACGCCCCCGGGC...GGTTCCG    1012
    ||||||||||||||||||||||||||:::  ||:::
316 yAlaGluAlaLeuCysProLeuValGlnThrProProAlaProThrSerG   333

1013 AGGTGGTCCCGTCATCAGCCGATGTACCAGCTCAC...AGGGCGTCG    1059
     |||||||||||||||||||  :::  |||
333 luMetValProValIleSerSerMetTyrProIleAlaLeuThrArgAla   349

1060 GAGGGCCACCCCGCTCCAACCACTCGCCCCAGGACCAGCGCCGTGGAGTA  1109
     |||  :::  ::|||  :::
350 Glu........MetSerAsnGlyAlaProGlnGluLeuGluArgLysSe    363

1110 CCTGCTGCTGTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGGAGGCGT    1159
     |||||||||        :::||:::
363 rIleLeuLeuProGluLysSer......ValProSerGluArgGlyLeuS   378

1160 CCCCGAGCAACAGCTGCCAAGACTCCACGGAGACAGAGCAACAACGAG    1209
     ||:::||||  :::|||||||||||  ||:::
378 erProAsnAsnSerGlyHisAspSerThrAspThrAspSerAsnHisGlu   394

1210 GAGCAGGCGCAGGGTCTTATCTACCTGACCAACCACCATCGCCGACGCGC  1259
     |||  :::  |||  :::
395 Glu...ArgGlnAsnHisIleTyrArgLnGlnAsnHisMetValLeuSerAr 410

1260 GCAACGCGTGTCG........CTCAAGGAGGAGCACCGGCCTACGACC    1300
     |||  ||:::
410 gAlaArgAsnGlyMetProLeuLeuLysGluValProArgSerTyrGluL   427

1301 TGCTGCGCGCGCCTCCGGAACTCGCCGAGACGCGCTCCGGTGGTCAGC    1350
     :::||||||||||||||||||||||||  :::|||
427 euLeuLysProProProIleCysProArgAspSerValLysValIleAsp   443

1351 ACCAGCGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCGGGTGCT    1400
     ::|||||  :::||||||||||||||||||||||||||||
```

us-08-711-417c-165.rspt

```
444 LysGluGlyGlyGluValMetAspValTyrArgCysAspHisCysArgValLe 460
1401 CTTCCTGGATCACGTCATGTACACCATCCACCATGGGCTGCCACGGCCTTCC 1450
          |||||||||:::||||||||||||||||||||||||||||||||||||||
460 uPheLeuAspTyrValMetPheThrIleHisMetGlyCysHisGlyPheA 477

1451 GTGATCCTTTGAGTGCAACATGTGCGGCTACCACAGCCAGGACACCGGTAC 1500
          |||||||||||||:||||||||||||||:::||:::||||||||||||||
477 rgAspProPheGluCysAsnMetCysGlyAspArgSerHisAspArgTyr 493

1501 GAGTTCTCCGCACATAACGGCGAGGGGAGCACCGC 1536
          |||||||::||:||:::||||||||||||||||
494 GluPheSerSerHisIleAlaArgGlyGlyGluHisArg 505
```

seq_name: sp_rodent:O08900

seq_documentation_block:
```
ID   O08900            PRELIMINARY;     PRT;    507 AA.
AC   O08900;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AIOLOS (FRAGMENT).
GN   ZNFN1A3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morgan B., Sun L., Avitahl N., Andrikopoulos K., Ikeda T.,
RA   Gonzales E., Wu P., Neben S., Georgopouolos K.;
RT   "Aiolos, a lymphoid restricted transcription factor that interacts
RT   with Ikaros to regulate lymphocyte differentiation.";
RL   EMBO J. 17:2004-2013(1997).
DR   EMBL; AF001293; AAB58795.1; -.
DR   MGD; MGI:1342542; Znfn1a3.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Zinc-finger.
FT   NON_TER    507     507
SQ   SEQUENCE    507 AA;    58048 MW;    58E459AE61D3D771 CRC64;
```

alignment_scores:
```
     Quality:  1302.00         Length:    526
     Ratio:    3.305           Gaps:       16
Percent Similarity: 74.905     Percent Identity: 52.281
```

alignment_block:
US-08-711-417C-165 x O08900

Align seg 1/1 to: O08900 from: 1 to: 507

```
  25 ATGTCTTTCTCATCAGGGAAGGAAGAAGCCCCCTGTAAGCGATACTCCAGA 74
          :::::::::|||       :::::|:::|||       :::::||:||||
   8 ValGluLeuLysSerThrGluGlnProLeuProThrGluSerProAs 24

  75 TGAG...GGCGATGAGCCCATGCCGATCCCGAGGACCTTCTCCACCACCT 121
          |      :::|||       ::|||||       :::|||||||||||
  24 pAlaLeuAsnAspTyrSerLeuProLysProHisGlulleGluAsnValA 41

 122 CGGGAGGACAGCAAAGTCCAAGAGTGACAGAGTCGTGGCCAGTAAT... 168
          :::||    :::       :::::::||||       |::::||||
  41 spSerArgGluAlaProAlaAsnGluAspGluAspGlyAlaGlyGluAspSer 57

 169 .......GTTAAAGTAGAGACTCAGATGATGAAGAAGAATGGGCGTGCCTG 212
          :::||       |||     :::||||       ::||||||
  58 MetLysValLysAspGluTyrSerAspArgAspGluAsnIleMetLysPr 74
```

```
213 TGAAATGAATGAAGAAGAATGTGCGGAGGATTTACGATGCTTGATGCCT 262
          |||   |||       |||||||       |||||||
 74 oGluProMetGly.....................AspAlaG 81

263 CGGGAGAGAAATGAATGGCTCCAACAGCCAAGGCAGCTCGGCCTTTG 312
          ::  :::::|||    |||:   |||:   |||:::
 81 luGluSerGluMetProTyrSerTyrAlaArgGlyTyrSerAspTyrGlu 97

313 TCG...............................GGAGTTGGAGGCATTCGACTTCCTAA 341
          :::      |||     |||       ::
 98 SerIleLysGluLeuArgHisValProTyrAspAsnSerArgProThrSe 114

342 CGGAAAACTAAAGTTGTGATATCTGTGGATCATTTGCATCGGCCCAATG 391
          :||||||:::||||:::||||||:       :||||||       |||
114 rGlyLysMetAsnCysAspValCysGlyLeuSerCysIleSerPheAsnV 131

392 TGCTCATGGTTCACAAAAGAGCCACACTGGAGAACGGCCCTTCCAGTGC 441
          |||||||||||||||||||||||||||||||||||:::||||||||||
131 aLeuMetValHisLysArgSerHisThrGlyGluArgProPheGlnCys 147

442 AATCAGTGCGGGCCTCATTCACCCAGAAGAGGCAACCTGCTCCGGCCACAT 491
          ||||||||||||||||||||||||||||||||||||||||||||||||||
148 AsnGlnCysGlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIl 164

492 CAAGCTGCCATTCCGGGAGGAGAGCCCTCAAATGCCACCTCTGCAACTACG 541
          |||||||||:::|||||||||||||||||||||||||||||||||||||
164 eLysLeuHisThrGlyGluLysProPheLysCysHisLeuCysAsnTyrA 181

542 CCTGCGCCCGAGGGACGCCCTCACTGGCCACTGAGGACGCACTCCGTT 591
          |||||||:::|||||||||||||||||||||||||||||||||||||
181 laCysGlnArgArgAspAlaLeuThrGlyHisLeuArgThrHisSerVal 197

592 GGTAAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAAC 641
          |||||:::|||     ::||||||:::|||||||||||||||||||:::
198 GluLysProTyrLysCysGluPheCysGlyArgSerTyrLysGlnArgSe 214

642 GTCTTTAGAGGAACATAAAGAGCCGCTGCCACAACTACTTGGAAAGCATGG 691
          :||||||||||||||||:::       :::||||||||||
214 rSerLeuGluGluHisLysGluArgCysArgAlaPheLeuGlnAsnProA 231

692 GCCTTCCGGGCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGT 741
          |||       :::||||||||||||       :::||||||||
231 spLeu...GlyAspAlaAlaSerVal..........GluAlaArgHisIle 243

742 GAAATGGCAGAAGACCTGCCAAGAATAGGATCCAGAGAGATCTCTCGTGCT 791
          :::::::       ||||||||||:::|||::|||||:::|||
244 Lys................AlaGluMetGlySerGluArgAlaLeuValLe 255

792 GGACAGACTAGCAGCAAGTAAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGA 841
          ||||||||||||||||||||||||||||||||||||:::|||||||
255 uAspArgLeuAlaSerAsnValAlaLysArgLysSerSerMetProGlnL 272

842 AATTCTTGGGGACAAG...GGCCGTGTCCGACACGCCCTACGACAGTGCC 888
          |||||:::|||       ::::|||||||||:::|||:::
272 ysPheIleGlyLuLysArgHisCysPheAspAlaAsnTyrAsnProGly 288

889 ...ACGTACGAGAAGGAGAACGAAGAAATGATGAAGTCCCACGTGATGGACCA 935
          ::       :::|||||||:::||       :::|||||||||
289 TyrMetTyrGluLysGluAsnGluMetMetGlnThrArgMetMetAspGl 305

936 AGCCATCAACAACGCCATCAACTACCTGGGGGCCGAGTCCTGCCCCGC 985
          |||||||||||||||||||||||||||||||||||:::|||||||||
305 nAlaIleAsnAsnAlaIleSerTyrLeuGlyAlaGluAlaArgProL 322

986 TGGTGCAGACGCCCCGGGC...GGTTCCAGGTGGTCCCGGTCATCAGC 1032
          ||||||||||||||||||       :::|||||||:::|||
322 euValGlnThrProProAlaProThrSerGluMetValProValIleSer 338

1033 CCGATGTACCAGCTGCAC..AGGCCTCGGAG.........GGCACCCC 1070
          ::||||       |||     :::|||       |||
339 SerValTyrProIleAlaLeuThrArgAlaAspMetProMetGlyAlaPr 355

1071 GCGCTCCAACCACTCGGCCCAGGACAGCCGCGTGGACTACCTGCTGCTGC 1120
```

Wed Aug 28 10:05:27 2002

US-08-711-417C-165 x Q96JP3

Align seg 1/1 to: Q96JP3   from: 1 to: 545

```
  1 CCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATG.....................  93
                    |||                 ::::          :::::
 52 CCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATG                        .
 52                ProSerCysSerLeuLeuGlnGlySerGlyAspSerSerLeuGluLysGl  30
 14 ProSerCysSerLeuLeuGlnGlySerGlyAspSerSerLeuGluLysGl

 94 ................CCGATCCCGAGGACCTCTCCACCACCTCCGGGAGGAC              130
                   ::::||||||||    ::: :::|
 30 uPheLeuGlyAlaProValGlyProSerValThrProAsn...SerG                   46

131 AGCAAAGTCCAAGAGTGACAGAGTCGTGGCCAGTAATGTTAAAGTAGAG                 180
    |:::|||||| |||   |||:::|||:|||||:||||:||||||||
 46 lnHisSerProSerProSerLeuSerAlaAsnSerIleLysValGlu                   62
```

seq_name: sp_human:Q96JP3

seq_documentation_block:
ID   Q96JP3   PRELIMINARY;   PRT;   545 AA.
AC   Q96JP3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   KIAA1782 PROTEIN (FRAGMENT).
GN   KIAA1782.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RX   MEDLINE=21245130; PubMed=11347906;
RA   Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XX.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 8:85-95(2001).
DR   EMBL; AB058685; BAB47411.1; -.
FT   NON_TER   1
SQ   SEQUENCE   545 AA; 59742 MW; 7A8539E5B8F9BD84 CRC64;

alignment_scores:
      Quality: 1296.50   Length: 552
        Ratio: 3.385   Gaps: 17
Percent Similarity: 69.384   Percent Identity: 50.181

alignment_block:

DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   EOS PROTEIN.
GN   ZNFN1A4 OR EOS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR;
RX   MEDLINE=99232954; PubMed=10218586;
RA   Homma Y., Kiyosawa H., Mori T., Oguri A., Nikaido T., Kanazawa K.,
RA   Tojo M., Takeda J., Tanno Y., Yokoya S., Kawabata I., Ikeda H.,
RA   Wanaka A.;
RT   "EOS: a novel member of the Ikaros gene family expressed predominantly
RT   in the developing nervous system.";
RL   FEBS Lett. 447:76-80(1999).
DR   EMBL; AB017615; BAA36213.1; -.
DR   HSSP; P15822; 1BBO.
DR   MGD; MGI:1343139; Znfn1a4.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 6.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   533 AA; 58167 MW; 7A5FF32C6FFDC372 CRC64;

alignment_scores:
       Quality: 1273.50   Length: 537
       Ratio: 3.387   Gaps: 18
       Percent Similarity: 70.019   Percent Identity: 50.652

alignment_block:
US-08-711-417C-165 x Q9Z2Z2

Align seg 1/1 to: Q9Z2Z2   from: 1   to: 533

    94 CCGATCCCCGAGGACCTTCCACCACCTCGGGAGGACAGCAAAGCTCCAA 143
        |||:| ||:::::::||||:|:||||::|||||:::||||||||||
    22 ProValGlyProSerValSerThrProAsn...SerGlnHisSerSerPr   37

   144 GAGTGACAGAGTCGTGGCCAGTAATGTTAAAGTAGACTCAGAGTGATG 193
        ||:| ||:::|||||||||||||::|||||::|||||||:||||||
    37 oSerArgSerLeuSerAlaAsnSerIleLysValGluMetTyrSerAspG 54

   194 AAGAGAATGGGCGTGCCTGTGAAATGAATGGGAAGAATGTGCGGAGGAT 243
        ||||||::|| ||| ||||||||||
    54 luGluSerSerArgLeu.......................LeuGlyProAsp 63

   244 TTACGAATGCTTGATGCCTCGGGAGAGAAAATG................... 276
        |||::|||||| ::| :::| |||
    64 GluArgLeuLeuAspLysAspSerValIleValGluAspSerLeuUse 80

   277 ...........AATGGCTCCCACAGGACCAAGGCAGCTCGG 307
             ::||||| ::|||||||||||||||
    80 rGluProLeuGlyTyrCysAspGlySerGlyProGluProHisSerPro. 96

   308 CTTTGTCCGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGT 357
        ::|||| |||||||||||||||||||||||||||||||||||||||||
    97 ....GlyGlyIleArgLeuProAsnGlyLysLeuLysCys 108

   358 GATATCTGTGGGATCAATTGCATCGGGCCCAATGTGCTCATGGTTCACAA 407
        |||:||||||:|::||||||||||||||||||||||||||||||||||||
   109 AspValCysGlyMetValCysIleGlyProAsnValLeuMetValHisLy 125

   408 AAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCT 457
        ||| |||||||||||||||||||||||||||||||||||||||||||||
   125 sArgSerHisThrGlyGluArgProPheHisCysAsnGlnCysGlyAlaS 142

   458 CATTCACCCAGAAGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGG 507
        |||||||||||||||||||||||||||||||||||||||||||||||||

```
                   |||||||||||||||||||||||||||||||||
142 erPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGly   158

508 GAGAAGCCCTCAAATGCCACCTGCAACTACGCCTGCCACCGGAGGGA        557
    ||||||||||||||  ::|||||||||||||||||||||||||||||
159 GluLysProPheLysCysProPheCysAsnTyrAlaCysArgArgArgAs     175

558 CGCCCTCACTGGCCACCTGAGGACGCACTCC........... GTTG          592
    ||||||||||||||||||||||||||||||           |||
175 pAlaLeuThrGlyHisLeuArgThrHisSerValSerSerProThrValG      192

593 GTAAACCTCACAAATGGGATATTGTGGCCGAAGCTATAAACAGCGAACG       642
    |||||||||:::||||||||||||||||||||||:::||||||||||
192 lyLysProTyrLysCysAsnTyrCysGlyArgSerTyrLysGlnInSer      208

643 TCTTTAGAGAACATAAAGAGCGCTGCCAACTACTTGGAAAGCATG..        690
    :::||||||||||||||||||||||||||||||||||:::|||
209 ThrLeuGluGluHisLysGluArgCysHisSerAsnTyrLeuGlnSerLeuSe   225

691 ...........GGCCTTCCGGCACACTGTACCCAGTCA                718
               || ||||||
225 rThrAspAlaGlnAlaLeuThrGlyGlnProGly.............       236

719 TTAAAGAGAAACTAAGCACAGTGAAATGCAGAAGAC..CTGTGCAAG         765
    :::||| ::|    |||   |||||||||||||||||
237 .....AspGluIleArgAspLeuGluMetValProAspSerMetLeuHis     251

766 ATAGGATCAGAGAGATCTCTCGTGCTGGACGACTAGCAAGTAATGTCGC      815
    :::||||||    :::||||||||||||||||||||||:::::::
252 ProSerThrGluArgProThrPheIleAspArgLeuAlaAsnSerLeuTh     268

816 CAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGACAAG.......      858
    :||||||||| |||  ||||||||||||:::|||
268 rLysArgLysArgSerThrProGlnLysPheValGlyLeuGlyGlnMetA     285

859 ......GGCCTGTCCGACACGCCCTACGAC...........AGTGCCACGTAC 894
          ||||||||||||||||         :::|||
285 rgPheSerLeuSerAspLeuProTyrAspValAsnAlaSerGlyGlyTyr     301

895 GAGAAGGAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAA     944
    |||||||    ||||:::||    :::::::||
302 GluLysAspValGluLeuValAlaAlaHisHisGlyLeuGluProGlyPheGl  318

945 CAACCCCATCAACTACTTGGGGGCCGAGTCCCTGCCGCCGTTGGTGCAGA     994
    :::::::::  :::||||||||||||||||||||||||::
318 yGlySerLeuAlaPheValGlyThrGluHisLeuArgProLeu..ArgL     334

995 CGCCCCCGGGCGGT.........TCCGAGGTGGTCCCGGTCATCAGCCCGATG 1038
    |||||||| ||||          :::||||||||
334 euProProThrAsnCysIleSerGluLeuThrProValIleSerSerVal    350

1039 TAC...........CAGCTGCACAGGGCCCTCGGAG......GGCACCCCC    1070
     |||           |||  |||||||||||   ||:::
351 TyrThrGlnMetGlnProIleProSerArgLeuGlnLeuProGlySerAr     367

1071 GCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGAGTACCTGCTGCTGC    1120
     |||||||||||||||||||||||::    :::       ||||
367 gGluAlaGlyGluGlyLysProGluAspLeuGly...AspGlyGlyProLeuL  383

1121 TCTCCAAGGCCAAGTTGGTGCCCTCGGAGGCGGAGGCCGTCCCCAGCAAC    1170
     || :::||::|    :::::::||  ||||||||||||||||||||
383 euTyrArgAlaArgGlySerLeuThrAspProGlyAlaSerProSerAsn    399

1171 AGCTGCCAAGACTCCACGGACACCGAGAGCAACAACGAGGAGCGCAG       1220
     :::||||||||||||||||||||||||||||||||||||||||||::
400 GlyCysGlnAspSerThrAspThrGluSerAsnHisGluAspArgIleGl     416

1221 CGGTCTT..............................A             1228
     |||
416 yGlyValGlyIleProSerSerGlySerProProGlnProProProThrI    433

1229 TCTACCTGACCAACCACCATCGCCCGACGGCGCCAACGCGGTGTCGCTCAAG  1278
     ||:  :::|||    |||
433 MetAspLeuSerAsnProGlnGlnProGlnLysAlaAla............ 13
```

```
433 leValValGlyArgHisHisSerProAlaTyrAla...............Lys 444

1279 GAGGAGCACCGCGCCTACGAC...CTGCTGCCGCGCCTCCGAGAACTC  1325
     |||: :::  :::          ||||||||||||||||||||||
445 GluAspProLysProGlnGlyLeuLeuLeuArgGlyThrProGlyProSe     461

1326 GCAGGACGCGCTCCGGTGGTCAGCACCCAGCGGGGAGCAGATGAAGGTGT 1375
     |::::::::|||||||||||:::  ||||||||||::::||||:
461 rLysGluValLeuArgValValGlyValGlyGluSerGlyValLysAlaP    478

1376 ACAAGTGCGAACACTGCCGCGGTGCTCTTCCTGGATCACGTCATGTACACC 1425
     ||||||||||||||||||||||||||||||||||||||||:::::::|||
478 heLysCysGluHisCysArgIleLeuPheLeuAspHisValMetPheThr    494

1426 ATCCACATGGGCTGCCACCGCTTCCGTGATCCTTTTGAGTGCAACATCTG 1475
     ||||||||||||||||||||||||||||||||||||||||||||||
495 IleHisMetGlyCysHisGlyPheArgAspProPheGluCysAsnIleCy    511

1476 CGGCTACCACCAGCCAGGACCCGTACGAGTTCTCGTCGCACATAACGCGAG 1525
     ||||||||||||||||||||||||||||||||||||||||||||||
511 sGlyTyrHisSerGlnAspArgTyrGluPheSerSerHisIleValArgG   528

1526 GGGAGCACCGC       1536
     |||||||||
528 lyGluHisLys 531
```

seq_name: sp_vertebrate:Q9PU55

seq_documentation_block;
ID Q9PU55   PRELIMINARY;   PRT;   494 AA.
AC Q9PU55;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LYMPHOID TRANSCRIPTION FACTOR.
GN AIOLOS.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RPL1; TISSUE=BURSA OF FABRICIUS, AND THYMUS;
RX MEDLINE=99438020; PubMed=10508239;
RA Liippo J.P., Mansikka A., Lassila O.;
RT "The evolutionarily conserved avian Aiolos gene encodes alternative
RT isoforms.";
RL Eur. J. Immunol. 29:2651-2657(1999).
DR EMBL; AJ005933; CAB56282.1; -.
DR HSSP; P15822; 1BBO.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR SMART; SM00355; ZnF_C2H2; 6.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE   494 AA; 55516 MW; 8D3AC945EDB31751 CRC64;

```
alignment_scores:
    Quality:   1260.00                Length:    539
    Ratio:        3.273                Gaps:       15
Percent Similarity:   71.429    Percent Identity:  49.351

alignment_block:
US-08-711-417C-165 x Q9PU55

Align seg 1/1  to: Q9PU55   from: 1  to: 494

25 ATGTCTTTCTCATCAGGGAAGGAAGGAAAGCCCCCCTGTAAGCCATACTCCAGA  74
   ||:||:::||||||||   :::|||
1 MetAspLeuSerAsnProGlnGlnProGlnLysAlaAla.......... 13
```

```
  75 TGAGGGCGATGAGCCATGCGATCCCCGAGGACCTCTCCACCACCTGG 124
     |||||::||| ::             ::
  14 .GluGlyGlnGluValLeu.........GluGluCysAspLeuSerL 26

 125 GAGGACAGCAAAGCTCCAAGAGTCACAGAGTCGTGGCCAGTAATGTTAAA 174
     |||||:: ::: ::||||::|||
  26 ysGlyGlnGluValGluGluValAlaAspAsnVal....GluGluLeuLys 40

 175 GTAGAGACTCAGAGTGATGAAGAGAAGAATGGGCGTGCCTGTGAAATGATGG 224
     ||||| ||| |||              :::::::
  41 GluHisSerGlnSerAsnGlyGluAlaGlyAspAspValLysValLysSe 57

 225 GGAAGAAATGTGCCGAGGATTACCAATGCTTGATGCCTCGGGAGAGAAAA 274
     ::||| ||:|||             |||::|:
  57 rGluTyrGlyAspArgGluGluSerAlaLeuAsnSer.....GluArgM 72

 275 TGAATGGCTCCCACAGGACCAAGGCAGCTCGGCT..........TTG 312
     ||:::: ::::::::||||| ||
  72 etGluAsnProGluGluSerGluLeuProTyrAlaTyrProArgGluTyr 88

 313 TCGGGAGTTGGAGCATTCGACTT................. 336
     |||
  89 SerGluTyrGluGlyIleIleLysGluArgHisLeuGlySerTyrArgAspSe 105

 337 ............CCTAACGGAAAACTAAAGTGTGATATCTGTGGATCATTT 376
     ::||||||||::: ||||||:::::: |
 105 rAlaArgProAlaSerGlyLysMetThrCysAspValCysGlyLeuAlaC 122

 377 GCATCGGGCCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACTGGAGAA 426
     |:||||: ||||||||||||||||||||||||||
 122 ysIleSerLeuAsnValLeuMetValHisLysSerHisThrGlyGlu 138

 427 CGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCATTCACCCAGAAAGGGCAA 476
     :::||||||:: :::||||||||:::::::||||
 139 ArgProPheGlnCysAsnGlnCysGlyAlaSerPheThrGlnLysGlyAs 155

 477 CCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAAGCCCTTCAAATGCC 526
     |||||||||||||||:||||||||||||| |||||||||
 155 nLeuLeuArgHisIleLysLeuHisThrGlyGluLysProPheLysCysH 172

 527 ACCTCTGCAACTACGCCTGCGGAGGGACGGCCCTCACTGGCCACCTG 576
     :::||||||:::||||||||||||||||||
 172 lsLeuCysSerTyrAlaCysGlyGlnArgArgAspAlaLeuThrGlyHisLeu 188

 577 AGGACGCACTCCGTTGGTAAACCTCACCAAATGTGGATAATTGTGGCCGAAG 626
     ||||||||||| ||||||:: ::||||||||
 189 ArgThrHisSerValGluLysProTyrLysCysGluPheCysGlyArgSe 205

 627 CTATAAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCGCTGCCACAACT 676
     ||||||||||::|||||||||||||||||| :::::::
 205 rTyrLysGlnArgSerLeuGluHisLysGluHisLysArgCysArgThrT 222

 677 ACTTGGAAAAGCATGGGC......CTTCCGGGCCACACTGTACCCAGTCATT 720
     :||:|||::||| |
 222 yrLeuGlnSerThrGlyGlyCysGluProAlaAsnVal............ 234

 721 AAAGAAGAAAACTAAGCACAGTGAAATGCAGAAGACCTGTCAAGAGATAGG 770
     ::||:::::||| :::||||
 235 ......GluAlaArgHisIleLys...........AlaGluMetGl 244

 771 ATCAGAGAGATCTCCGTGCTGGACAGACTAGCAAGTAATGTCGCCAAAC 820
     |:::||||||||||||||||||||||||||||||
 244 yThrGluArgAlaLeuValLeuAspArgLeuAlaSerAsnValAlaLysA 261

 821 GTAAGAGCTCATAGCCTCAGAAATTTCTTGGGACAAGGGCCTGTCC... 867
     |||||||||||||: :::|||
 261 rgLysSerMetProGlnIlysPheIleGlyGluLysGlyArgHisSerPhe 277

 868 GACACCGCCCTACGACAGTGCC...ACGTACGAAGAGGAGAACGAAATGAT 914
     |||::: |||:: :||||||||||||||||||||
 278 AspValAsnTyrAsnSerSerPheValTyrGluLysGluGlyAspIleVa 294
```

```
 915 GAAGTCCCACGTGATGGACCAAGCCATCAACAACACGCCATCAACTACCTGG 964
     :::::::||||||:: ||| ||||||||||||||::::||:::||
 294 lGlnGlyArgMetMetAspGlnAlaIleAsnAsnAlaIleSerPheLeuG 311

 965 GGGCCGAGTCCCTGCCCCGCCTGGTGCAGAGCGCCCCCGGGC...GGTCC 1011
     |||||||||||||||||||| ::: :::
 311 lyAlaGluAlaLeuArgProLeuValGlnThrProProAlaProAlaPro 327

1012 GAGGTGGTTCCGGTCATCAGCCACCGATGTACCAGCTGCACAGGCGGCTCCGA 1061
     ||||::||||||||||||:: :::
 328 GluMetValProValIleSerProLeuTyrProIleAlaLeuProArgAl 344

1062 GGGCACCCCGGCTCCAACCACTCGGCCCAGGACAGCCCGTGGAGTACC 1111
     ::||:::::::::||
 344 aAspValProAsnGlyGlyAsp............ 351

1112 TGCTGGTCGTCTCCAAGGCCAGTTGTGCCCTCGGAGCGCGAGGGCGTCC 1161
     ::||||: :: |||||
 352 ..........GluLysSerHisValProLeuArgGluArgAlaLeuSer 364

1162 CCGAGCAACAGCTGCCAAGACTCCACGGACACCAGGAGCAACAACAAGCAGGA 1211
     |||::||||| :::||||||||||||||:::|||
 365 ProAsnAsnSerGlyHisAspSerThrAspSerAsnHisGluGl 381

1212 GCAG.............................................CGCA 1219
     :::
 381 uArgProAsnProProPheProGlnGlyProValLeuProValProArgA 398

1220 GCGGTCTTATCTACCTGACCAACCACCATCGCCCGACGCCGCCAACCGCTG 1269
     ::||||| .................................Pro 401
 398 snGlyLeu.

1270 TCGCTCAAGGAGGAGCACCGCGCCTACGACCTGCTCGCCGCCGCCTCCGA 1319
     :::||||||| :::|||:::::::::::::||||
 402 AlaLeuLysAspPheProArgProPheAspSerIleIleLysProProAlaMe 418

1320 GAACTGCAGAGACGCGCTCCGCGTGGTCAGCAGCACCAGCGGGGAGCAGATGA 1369
     ::||||||:::||||| :::||||||
 418 tCysProArgAspAlaPheLysValIleAsnLysAspGlyGluAlaIleG 435

1370 AGGTGTACAAGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGTCATG 1419
     :::|||||||::||||||||||||||||||||||||||:::||
 435 lyAlaTyrArgCysAspHisCysArgValLeuPheLeuAspTyrValMet 451

1420 TACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAA 1469
     ::|||||||||||||||||||||||||||||||||||||||||||
 452 PheThrIleHisMetGlyCysHisGlyPheArgAspProPheGluCysAs 468

1470 CATGTCGCGCTACCACCAGGACGACCGGTACGAGTTCTCGTCGCACATAA 1519
     ||||||||||||||||||||||||||||||||||||||||||||||
 468 nValCysGlyTyrArgSerHisAspArgTyrGluPheSerSerHisIleA 485

1520 CGCGAGGGGAGCACCGC 1536
     ::||||||||||||
 485 laArgGlyGluHisArg 490
```

seq_name: sp_human:Q9H2S9

```
RT   "Eos and Pegasus, Two Members of the Ikaros Family of Proteins with
RT   Distinct DNA Binding Activities.";
RL   J. Biol. Chem. 275:38347-38354(2000).
DR   EMBL; AF230809; AAG39221.1; -.
DR   HSSP; P15822; 1BBO.
DR   InterPro; IPR000822; Znf-C2H2.
DR   Pfam; PF00096; zf-C2H2; 6.
DR   SMART; SM00355; ZnF_C2H2; 6.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; zinc-finger.
SQ   SEQUENCE   483 AA;  53219 MW;  C30B0270709C34F9 CRC64;

alignment_scores:
         Quality:   1252.50               Length:    500
           Ratio:      3.528                 Gaps:     15
Percent Similarity:  71.000      Percent Identity:   52.400

alignment_block:
US-08-711-417C-165 x Q9H2S9

Align seg 1/1   to: Q9H2S9   from: 1   to: 483
```

```
188  ...........AspGluIleArgAspLeuGluMetValProAspSerMe 200
        :::|||                   ::                  |||
759  GTGCAAGATAGGATCAGAGAGATCTTCGTGCTGACAGACTAGCAAGTA 808
        ::|||||||||::     ::|||||||||||:::
200  tLeuHisSerSerGluArgAlaThrPheIleAspArgLeuAlaAsnS 217

809  ATGTCGCCAAACGTAGGAGCTCTATGCTCAGAAATTCTTGGGACAAG 858
        ||||||||::||:  ||||||||::::||||::||:||::||||||||
217  erLeuThrLysArgLysArgSerThrProGlnLysPheValGlyLuLys 233

859  ...........GGCCTGTCCGACACGCCCTACGAC......AGTCCCAC 890
        ::|||||||||   |||||||||         ||||
234  GlnMetArgPheSerLeuSerAspLeuProTyrAspValAsnSerGlyGl 250

891  GTACGAGAAGGAGAACGAAATGATGAAGTCCCACGTGATGACCAAGCCA 940
        |||||||||::   |||:::::   |||        ::::
250  yTyrGluLysAspValGlyLeuValGlyHisSerLeuGluProGlyP 267

941  TCAACAACGCCATCAACTACCTGGGGCCGAGTCCCTGCGCCGTGGTG 990
        :::||||||||||||||||||||||||||||||||||||||||||||
267  heGlySerSerLeuAlaPheValGlyAlaGluHisLeuArgProLeu... 282

991  CAGAGCGCCCCGGCGGGT......TCCGAGGTGGTCCCGGTCATCAGCCC 1034
        ::::|||||||||||||      |||||||||   :||||||||||||
283  ArgLeuProProThrHisCysIleSerGluLeuThrProValIleSerSe 299

1035 GATGTAC........CAGCTGCACAGGCGCTCGGAG.....GGCA 1066
        ::|||       |||    |||||||||||||:||    |||
299  rValTyrThrGlnMetGlnProLeuProGlyArgLeuGluLeuProGlyS 316

1067 CCCCGGCTCCAACCACTCGGCCCAGACAGCCGCGTGGAGTACCTGCTG 1116
        ::   |||||||||||||||||||   :|:|||   :|||||||||||
316  erArgGluAlaGlyLeuGlyProGluAspLeuAla...AspGlyGlyPro 331

1117 CTGCTCTCCAAGGCCAAGTTGGTCCCCTCGGACGCGCAGGCGTCCCCGAG 1166
        ||||||         |||||||||||||||||||||||||||
332  LeuLeuTyrArgProArgGlyProLeuThrAspProGlyAlaSerProSe 348

1167 CAACAGTGCCAAGACTCCACGGACACCGAGAGCAACAGGAGCAGC 1216
        ::::|||||||||    |||||||||||||||||||||||||||
348  rAsnGlyCysGlnAspSerThrAspThrGluSerAsnHisGluAspArgv 365

1217 GCAGGCGGTCTTATCTACCTG 1236
        ::|||||   ||
365  aIAlaGlyValValSerLeuProGlnGlyProProThrGlnProProPro 381

1237 ACCAACCACATCGCCGACGCGGGCCAACGCGTCGCTCAAGGAGGAGCA 1286
        |||       :::|||||      :::||||||||||||||||||||||
382  ThrIleValGlyValArgHisSerProAlaTyrAlaLysGluAspProLy 398

1287 CCCCGGCCTACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGC 1336
        :::|||||||||||||||||||||||||||||||||||||||||||||
398  sProGlnGluGlyLeuLeuArgGlyThrProGlyProSerLysGluValL 415

1337 TCCGCGTGGTCAGGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAA 1386
        :::|||||||     |||||||||||||||||||||||||||||||||
415  euArgValGlyLeuValSerSerGluProValAlaPheLysLysCysGlu 431

1387 CACTGCCGGGTGCTCTTCCTGATCACGTCATGTACCACCATCCACATGG 1436
        ||||||||||||||||||||||||::|||||||||||||||||||||
432  HisCysArgIleLeuPheLeuAspHisValMetPheThrIleHisMetGl 448

1437 CTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTCGGCTACCACA 1486
        :::|||||||||||||||||||||||||||||||||||||||||||
448  yCysHisGlyPheArgAspProPheGluCysAsnIleCysGlyTyrHisS 465

1487 GCCAGGACCGTACGAGTTCTCGCACATAACGCGAGGGAGCACCGC 1536
        |||||||||||||||||||||||||||||||||||||||
465  erGlnAspArgTyrGluPheSerHisIleValArgGlyHisHisLys 481
```

```
187  AGTGATGAGAGAATGGG...........................CGTGCCTGTGA 215
        |||||||||||:::::::                          ||  ||
3    SerAspGluSerSerArgLeuLeuGlyProAspGluArgLeuLeuGl 19

216  AATGAATGGGGAGGAAGAATGTGCGAGGAT..............TTACGAA 250
        |||  :::|||||||||||||::||||||||              ||
19   uLysAspAspSerValIleValGluValAspSerLeuSerGluProLeuGlyT 36

251  TGCTTGATGCCTCGGGAGAGAAAATGAATGAGTCCCCAGGGACCAAGGC 300
        ||||||||:|||||   ||| |||     ::::::
36   yrCysAspGlySerGlyProLysProLysProHisSerPro........ 47

301  AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
        |||||||||||||||||||||||||||||||||||||||||||
48   ....GlyGlyIleLeuArgLeuProAsnGlyLysLysLe 57

351  AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
        ||||||||:::||||||:|:||||||||||||   :||||||||||||
57   uLysCysAspValCysGlyMetValCysIleGlyProAsnValLeuMetV 74

401  TTCACAAAGAAGCCACACTGGAGACACGGCCCTTCCAGTGCAATCAGTGC 450
        |||||||||||||||||||||||||||||||:||||||||||:||||||
74   alHisLysArgSerHisThrGlyGluArgProHisCysAsnGlnCys 90

451  GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCCACATCAAGGTGCA 500
        ||||||||||||||||||||||||||:||||||||||||||||||||||
91   GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLysLeuHi 107

501  TTCCGGGAGAAGCCCTTCAAATGCCACTACGCCTGCCGCC 550
        |||||||||||||||   :::|||||||||||||||||||
107  sSerGlyGluLysProPheLysCysProPheCysAsnTyrAlaCysArgA 124

551  GGAGGACGGCCCTCACTGGCCACCTGAGGACGCACTCC............ 588
        ||||||||||||||||||||||||||||||||||||
124  rgArgAspAlaLeuThrGlyHisLeuArgHisThrHisSerValSerSerPro 140

589  ..GTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAAACA 635
        |||||||||||||||||||||||||:||||||||||||||||||||
141  ThrValGlyLysProTyrLysCysAsnTyrCysGlyArgSerTyrLysGl 157

636  GCGAACGTCTTTAGAGGAACATAAAGAGGCTGCCACAACTACTTGGAAA 685
        |:::::||||||||||||||||||||||||||||||||||||||:::
157  nGlnSerThrLeuGluGluHisLysLysGluArgCysHisAsnTyrLeuGlnS 174

686  GCATG............................GGCCTTCCGGCCACACTGTAC 711
        |:::|                           ||| ||||||
174  erLeuSerThrGluAlaGlnAlaLeuAlaGlyGlnLnProGly......... 187

712  CCAGTCATTAAAGAAGAAGAAACTAAGCACACAGTGAAATGGCAGAAGAC...CT 758
```

us-08-711-417c-165.rspt

Wed Aug 28 10:05:27 2002